STUDIES ON NEW RESTRICTION ENZYMES AND MODIFICATION ENZYMES FROM BACTERIA(Dissertation_全文)

AUTHOR(S):

Kita, Keiko

STUDIES ON NEW RESTRICTION ENZYMES AND MODIFICATION ENZYMES FROM BACTERIA

KEIKO KITA

1989

STUDIES ON
NEW RESTRICTION ENZYMES
AND
MODIFICATION ENZYMES FROM BACTERIA

KEIKO KITA

1990

CONTENTS

ABBREVIATIONS

AdoMet	S-adenosyl-L-methionine
bp	base pair
BSA	bovine serum albumin
ds	double-stranded
EDTA	ethylenediaminetetraacetic acid
HPLC	high-performance liquid chromatography
IPTG	isopropyl-1-thio-β-D-galactopyranoside
kb	kilobase
kDa	kilodalton
Km	Michaelis constant
PAGE	polyacrylamide gel electrophoresis
R*Fok*I etc.	*Fok*I restriction endonuclease etc.
M*Fok*I etc.	*Fok*I methylase etc.
SDS	sodium dodecyl sulfate
ss	single-stranded
Tm	midpoint melting temperature
Vmax	rate of the enzyme-catalysed reaction at an infinite concentration of the substrate

INTRODUCTION

Restriction enzymes are strain-specific endonucleases that enable bacteria to recognize and destroy foreign DNA rapidly and which cause double-stranded scission at a limited number of sites of the DNA. In addition to having restriction activity, each bacterial strain with at least one such endonuclease has a specific DNA methylase that transfers methyl groups from AdoMet to specific residues in the DNA. DNA modified in this way is resistant to cleavage by co-existing restriction enzyme. In this way, the DNA of the cell is protected from its own restriction enzymes.

The phenomenon of host-controlled restriction-modification has been known for a long time. In 1968, Meselson and Yuan first isolated and characterized a DNA restriction enzyme from Escherichia coli strain K (1). The enzyme, EcoK, recognizes a specific sequence but cleaves randomly, so it has not been used in practical applications. In 1970, Smith and Wilcox isolated a restriction enzyme, HincII, from Haemophilus influenzae Rc (2). The enzyme recognizes GTPyPuAC and cleaves between Py and Pu. This was the first characterization of a type II restriction enzyme. By use of this enzyme, it became possible to cleave DNA at a specific site. Since then, screening for restriction enzymes has continued and enzymes with different specificities have been isolated from a variety of organisms. So far, about 1,000 enzymes have been discovered (3).

Restriction enzymes can be grouped into three classes, I, II, and III, on the basis of subunit structure, cofactor requirements, substrate specificity, and other features (4). Type I enzymes (EC 3.1.21.3) have three subunits, have both methylase and endonuclease activities in the same protein assembly, require Mg^{2+}, ATP, and AdoMet for cleavage, and cleave almost at random at a great distance from the recognition site. Type II enzymes (EC 3.1.21.4) are usually homodimers, co-existing with a separate methylase protein, requiring Mg^{2+} for endonuclease but not methylase activities, and cleaving within the recognition site or a few nucleotides away. Type III enzymes (EC 3.1.21.5) have two subunits, have both methylase and endonuclease activities in the same protein assembly, require Mg^{2+} and ATP for cleavage (AdoMet stimulates cleavage but is not required), and cleave tens of nucleotides away from the recognition site. Recently, two restriction enzymes were found that do not seem to fit into any of these classes (5). Both enzymes are active in the presence of Mg^{2+} alone, with no requirement for ATP. However, both are stimulated by low concentrations (10^{-7}-10^{-4} M) of AdoMet. They have been tentatively classified as type IV.

Usually, type II restriction-modification enzymes are used for gene manipulation, and their use has enabled revolutionary advances to be made in biological research. Type II enzymes recognize 4-8 nucleotides with two-fold rotationary symmetry and cleave within the recognition sequence. At present, almost 1,000 type II enzymes with a total of 110 different specificities have

been discovered (3). More than twenty restriction enzymes of type II, for which the cleavage site is shifted a considerable but precise distance (up to 20 nucleotides) from the recognition site, have been described (6). Such enzymes are referred as 'shifter' or 'type IIS' enzymes. More are being found each year. The list of known restriction enzymes contains many derived from different organisms but with identical sequence specificity. These are called isoschizomers. Some have a different sensitivity to methylation in the recognition sequence, and some cleave at a different position in the recognition sequence (7).

There are two DNA methyltransferases; one (EC 2.1.1.72) modifies a specific adenine at the N^6 position within the target sequence and the other (EC 2.1.1.73) modifies cytosine at C^5, or less commonly, N^4, again, within the target sequence (8,9). Not only restriction enzymes but also methyltransferase is used for gene manipulation as follows. When a DNA fragment is cloned with use of a linker, the DNA to be inserted is methylated before ligation with the linker to protect it from digestion by restriction enzymes. New DNA cleavage specificities can be created by the combination of restriction enzymes and methylases (10).

Restriction enzymes are found mainly in bacteria; some are found in algae and viruses. Some of these sources are pathogenic, and care must be taken in their handling. The production of the enzymes is different in different species and some enzymes are produced in amounts too small for commercial use or for

biochemical research. To overcome such difficulties, cloning of the gene for the restriction-modification system in E. coli cells has been tried. In 1978, a first report of the cloning of a complete restriction-modification system of Haemophilus haemolyticus in E. coli appeared (11). To date, more than 60 restriction-modification systems have been cloned completely and a further 40 have been cloned in part (12). Analysis of the cloned system by gene and protein sequencing is proceeding rapidly. Well-conserved blocks of amino acid sequences have been determined in the groups of C- (13) or A-methylases (14,15), which recognize different DNA sequences even in the same group. Genetically engineered overproducing strains are suitable for use in the production of enzymes. Some restriction endonucleases have been crystallized from such strains and analysed by X-ray diffraction (16,17). The analysis is expected to throw light on enzyme structure and catalysis, and to clarify the apparently numerous ways in which sequence-specific protein-DNA recognition occurs.

In this thesis, I describe the studies on the restriction-modification enzymes. In chapter I, I describe the rapid method for determination of the cleavage site of restriction enzyme. In chapter II, I describe the interaction of the restriction enzyme with its substrate. The purification and characterization of a new restriction enzyme are described in chapter III. In chapter IV to VII, the studies on the restriction-modification system of Flavobacterium okeanokoites are described. The isolation and the nucleotide sequence of the gene are described in chapter IV, as is

the overexpression of FokI endonuclease in E. coli cells in chapter V. FokI methylase is also purified and characterized (chapter VI). Finally, I discuss the role of amino acid segment of FokI methylase in the methylation of the recognition sequence.

CHAPTER I

Determination of the cleavage site of restriction enzyme, AccII, using synthetic oligonucleotide

The site-specific endodeoxyribonuclease, AccII, is known to be an isoschizomer of FnuDII, but its cleavage site has not yet been determined (18). This chapter deals with determination of the cleavage site of AccII using the self-complementary decanucleotide, d(CTACGCGTAG), as substrate. The cleavage site of this enzyme was deduced to be 5'-CG↓CG-3'.

AccII was purified as follows. About 150 g cells were disrupted by sonication and the debris was removed by centrifugation (10^5 x g for 1 h). The fraction precipitated with 50% (w/v) ammonium sulfate was collected and suspended in KP buffer (10 mM potassium phosphate, pH 7.5, 10 mM 2-mercaptoethanol and 5% glycerol). After dialysis against KP buffer, 0.2% (v/v) Polymin P was added to remove nucleic acids. The enzyme solution was then applied to phosphocellulose (Whatman P11) column and eluted with a linear gradient of 0-1.0 M KCl in KP buffer. The AccII active fractions (0.55-0.63 M KCl) were pooled, applied to a heparin-Sepharose (Pharmacia CL-6B) column, and eluted with a linear gradient of 0-1.5 M KCl in KP buffer. The active fractions (0.12-0.32 M KCl) from this column were pooled, applied to an Affi-Gel Blue agarose (Bio-Rad) column, and eluted with a linear

gradient of 0-1.5 M KCl in KP buffer. Finally, the active fractions (0.07-0.21 M KCl) were pooled and applied to a hydroxylapatite (Clarkson) column. After elution with a linear gradient of 0-2.0 M KCl in KP buffer, the active fractions (0.20-0.62 M KCl) were pooled, diluted with an equal volume of glycerol and stored at -20°C. Enzyme activity was measured in a reaction mixture (50 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 60 mM NaCl, 0.01% BSA, and 1 µg of λ-DNA. One unit is defined as that amount of enzyme required to digest completely 1 µg of λ-DNA in 60 min at 37°C. No non-specific nuclease activity, as determined by agarose gel electrophoresis, was detected after incubation of 1 µg of λ-DNA with 37.5 units AccII for 24 h.

Synthesis of the decanucleotide d(CTACGCGTAG) was carried out by a solid phase method (19). The 5'-terminal hydroxyl group was

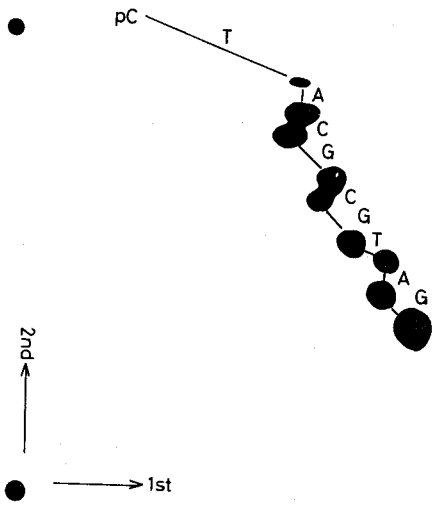

Figure 1. DNA sequence of the decanucleotide, ^{32}P-CTACGCGTAG.

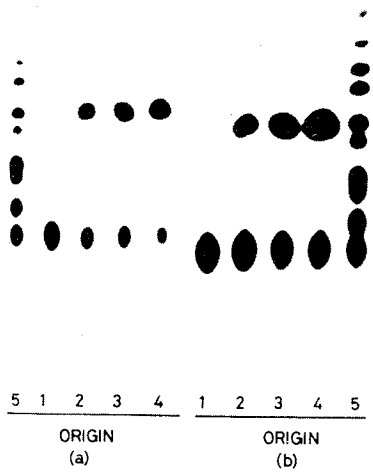

Figure 2. Autoradiograms of oligonucleotides produced on AccII and FnuDII digestion. See the text for details of AccII(a) and FnuDII (b) digestion. Incubation times were as follows: 1, 0 min; 2, 15 min; 3, 30 min; 4, 90 min. A snake venom phosphodiesterase digest, 5, was used as authentic markers.

then phosphorylated with [γ-^{32}P]ATP and polynucleotide kinase. T4 polynucleotide kinase (4 units) was added to a reaction mixture (5 μl) containing 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 100 pmol [γ-^{32}P]ATP (10^6cpm), and 100 pmol d(CTACGCGTAG), and incubated at 37°C for 1 h. The sequencing was carried out by the standard method (20) (Fig. 1).

Digestion of ^{32}P-CTACGCGTAG was performed with AccII and FnuDII as follows. AccII (37.5 units) was added to a reaction mixture (50 μl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 60 mM NaCl and 25 pmol ^{32}P-CTACGCGTAG, and incubated at 15°C. FnuDII (20 units), purchased from New England Biolabs., was added to a reaction mixture (20 μl) containing, 6 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 6 mM NaCl and 25 pmol ^{32}P-CTACGCGTAG, and incubated at 15°C. In each case, the reaction was stopped by heating at 100°C for 2 min. The oligonucleotides obtained were then separated by

homochromatography on a DEAE-cellulose thin-layer plate using Homomixture II. Autoradiograms are shown in Fig. 2. The pentanucleotide, ^{32}P-CTACG, was detected as the product of FnuDII digestion. This indicates that FnuDII reacts with the synthetic decanucleotide in the same way as it does with DNA (21). Similarly, the pentanucleotide, ^{32}P-CTACG, was detected as the product of AccII digestion. From these results, it is concluded that AccII cleaves DNA at the sites indicated by the arrow, CG↓CG, and at the position as FnuDII.

Usually, the cleavage site of a restriction enzyme is determined using plasmid- or viral DNA as substrate. It may be concluded that synthetic oligonucleotide can serve as a substrate for the determination of the cleavage sites of restriction enzymes.

CHAPTER II

Interaction of the restriction endonuclease ScaI with its substrates

Many type II restriction enzymes have been isolated from bacteria, and their specificities characterized (22). They recognize a specific nucleotide sequence of 4 to 8 bp in double-stranded DNA and cleave within or near the recognition sequence. These enzymes are indispensable for genetic engineering, but the mechanism of their recognition and cleavage is unknown. Recently, a simple and economical method (19) for the synthesis of oligodeoxyribonucleotides was developed, and synthetic DNA is now readily available. Using chemically synthesized DNA containing base analogues, some researchers have made suggestions about the possible mode of recognition and cleavage of specific nucleotide sequences by restriction enzymes (23-25). Kinetic parameters of EcoRI for various natural and synthetic substrates have been analysed in detail (26,27). In this chapter I describe the interaction of ScaI with synthetic oligonucleotides of various lengths and also with native DNA.

MATERIALS AND METHODS

DNA and enzymes λ phage DNA was prepared from E. coli K 12 W3350 (λcI857 S7) by the procedure of Thomas and Davis (28). Plasmid pBR322 DNA was isolated from a transformed E. coli C600, as described by Guerry et al. (29). The structures of the oligonucleotides used here are given in Table I. They were synthesized using a solid-phase method (19). The 5'-terminal hydroxyl group was phosphorylated with [γ-^{32}P]ATP and polynucleotide kinase, and labeled oligonucleotides were separated on a Sephadex G-50 (Pharmacia) column. Another restriction endonuclease, PvuII, was prepared from Proteus vulgaris, alkaline phosphatase from E. coli C75 by the procedure of Nakata et al. (30), and polynucleotide kinase from E. coli B infected with phage T4amN82 by the method of Panet et al. (31).

Bacterial strain and culture conditions Streptomyces caespitosus was used. Cells were aerobically incubated at 26°C in L-broth (1% bacto-tryptone, 0.5% yeast extract, 0.1% glucose, and 0.5% NaCl, pH 7.2) and harvested by centrifugation when the late logarithmic phase was reached. The yield of cells per liter culture was about 17 g (wet weight).

Purification of ScaI ScaI was purified by methods described elsewhere (32), with slight modifications. About 230 g of cells were treated with 580 mg of lysozyme and disrupted by sonication. Debris was removed by centrifugation (10^5 x g for 1 h). The supernatant was treated with 2% (w/v) streptomycinsulfate

to remove nucleic acids. The fraction precipitated with 40-80% (w/v) ammonium sulfate was collected and suspended in KP buffer (10 mM potassium phosphate, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerol). After dialysis against the same buffer, the enzyme solution was put on a DEAE-cellulose (Whatman DE52) column (6 x 10 cm), and eluted with a linear gradient of 0-0.6 M KCl in KP buffer. The active fractions with the ScaI (0.14-0.28 M KCl) were pooled, dialyzed, put on a DEAE-cellulose column (4 x 8 cm), and eluted with a linear gradient of 0-0.6 M KCl in KP buffer. The active fractions (0.18-0.25 M KCl) were pooled, dialyzed, put on a Affi-Gel Blue agarose (Bio-Rad) column (1.5 x 16 cm), and eluted with a linear gradient of 0-1.0 M KCl in KP buffer. The active fractions (0.58-1.0 M KCl) were pooled, diluted with an equal volume of glycerol, and stored at -20°C.

Assay of ScaI endonuclease Enzyme activity was measured in a reaction mixture (50 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 125 mM NaCl, 7 mM 2-mercaptoethanol, 0.01% BSA, and 1 µg of DNA. One unit was defined as the amount of enzyme required to digest 1 µg of λ-DNA completely in 60 min at 37°C.

Phosphodiester bond cleavage with ^{32}P-labeled oligonucleotides as substrates was measured by following the release of ^{32}P-labeled product. The reaction mixture (50 µl) contained 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 125 mM NaCl, 7 mM 2-mercaptoethanol, 0.125-7 µM oligonucleotide, and 75 units of ScaI. Incubation was at 15°C. Portions of the reaction mixture were taken, the reaction was stopped by boiling for 2 min in the

presence of 50 mM EDTA, and the portions were spotted on a DEAE-cellulose thin-layer plate, which was then developed in Homomixture II (20). Spots were located by autoradiography. The spots corresponding to the product and to the uncleaved substrate were scraped from this plastic-backed plate, and ^{32}P was determined by Cherenkov counting. The extent of reaction was calculated by:

phosphodiester bonds cleaved (pmol) =

$$\frac{\text{cpm (product)}}{\text{[cpm (product) + cpm (uncleaved substrate)]}} \times \text{total amount of substrate (pmol)}$$

The total amount of substrate was calculated from its absorbance at 260 nm.

Cleavage of phosphodiester bonds with pBR322 DNA as the substrate was measured as follows. pBR322 DNA was cleaved by PvuII to make linear DNA. The 5'-hydroxyl terminal was dephosphorylated with alkaline phosphatase, and phosphorylated with [γ-^{32}P]ATP and polynucleotide kinase using the method of Maxam and Gilbert (33). ^{32}P-Labeled DNA fragments were separated from unincorporated [γ-^{32}P]ATP by chromatography on a Sephadex G-50 column. The rate of phosphodiester bond cleavage was measured by following the release of ^{32}P-labeled fragments 1780 or 2583 bp long. The reaction mixture (150 µl) contained 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 125 mM NaCl, 7 mM 2-mercaptoethanol, 2.5-75 nM DNA, and 75 units of ScaI. Incubation was at 15°C. Portions (15 µl) were taken and the reaction was stopped by the addition of 5 µl of a solution containing 200 mM NaOH, 200 mM EDTA, 20% sucrose,

and 1% bromocresol green. The DNA fragments were analysed by alkaline agarose gel electrophoresis on a 0.7% agarose slab gel, by slight modification of McDonell's method (34). Here, gel was prepared in a neutral solution (50 mM NaCl and 1 mM EDTA), and equilibrated in alkaline electrophoresis buffer (30 mM NaOH and 1 mM EDTA) before the run. The DNA bands were cut out of the gel and melted in 0.4 ml of H_2O; the amount of ^{32}P it contained was measured after addition of 6 ml of Aquazol-2 scintillant. The extent of the reaction was estimated by:

phosphodiester bonds cleaved (pmol) =

$$\frac{[\text{cpm (1780 bp)} + \text{cpm (2583 bp)}]}{[\text{cpm (1780 bp)} + \text{cpm (2583 bp)} + \text{cpm (4363 bp)}]} \times \text{total amount of DNA (pmol)}$$

The amount of DNA was calculated from its absorbance at 260 nm.

Confirmation of ScaI cleavage site Cleavage site was checked using d(CGCAGTACTGCG) as the substrate by the method described in chapter I. ScaI (150 units) was added to a reaction mixture (50 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 125 mM NaCl, and 0.6 µM 5'-^{32}P-labeled oligonucleotides, and the solution was incubated at 15°C for 15 to 120 min. The resulting oligonucleotides were then separated by homochromatography (20) on a DEAE-cellulose thin-layer plate using Homomixture II, and detected by autoradiography.

Chemicals DEAE-cellulose thin-layer plates (CEL 300 DEAE/HR-2/15) from Masherey-Nagel Co. and [γ-^{32}P]ATP from Amersham Inc. All other reagents were of the highest purity available.

RESULTS

Confirmation of ScaI cleavage site Determination of the cleavage site of ScaI was reported elsewhere (32). The cleavage site was determined using an M13mp18 derivative containing an AGTACT sequence, which was constructed by inserting a synthetic octanucleotide d(AAGTACTT). I tried to determine the cleavage site directly using the octanucleotide. However, it was not cleaved efficiently, especially when its concentration was low. The cleavage site of ScaI was checked using the synthetic dodecanucleotide d(CGCAGTACTGCG). The hexanucleotide ^{32}P-CGCAGT was found to be the product of ScaI digestion (Fig. 1). This shows that ScaI cleaves the substrate at CGCAGT↓ACTGCG, where indicated by the arrow.

Interaction of ScaI with its substrate In the course of cleavage site confirmation, the cleavage rate seemed to be dependent on the length of substrate. For this reason, I investigated the affinity of ScaI to oligonucleotides of other lengths. I synthesized self-complementary oligonucleotides that contained the ScaI recognition site at the middle of the oligonucleotide sequence. The nucleotide sequence of each oligonucleotide as identified by two-dimensional homochromatography was in complete agreement with the assigned structure (data not shown). Tm values for each oligonucleotide in the reaction mixture were calculated using the equation of Davis et al. (35), and are listed in Table I. Tm for the octanucleotide

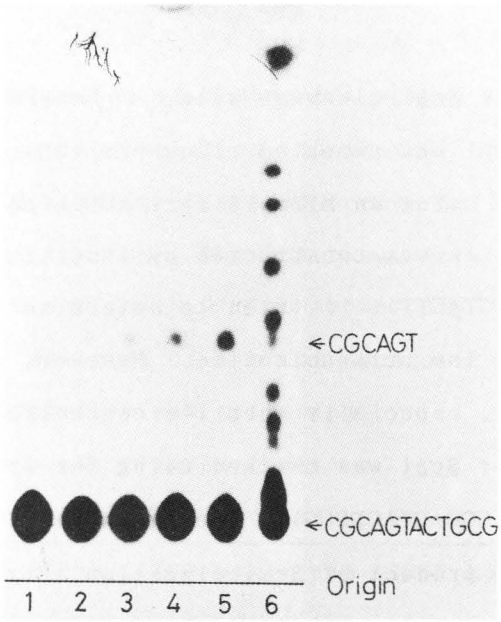

Figure 1. Identification of oligonucleotides produced from synthetic dodecanucleotide by digestion with ScaI. The dodecanucleotide was digested with ScaI for different times (lane 1, 0 min; lane 2, 15 min; lane 3, 30 min; lane 4, 60 min; lane 5, 120 min). Lane 6 is of authentic markers obtained by digestion with venom phosphodiesterase.

was 14.3°C, so I measured the rate at which ScaI cleaved the oligonucleotides and pBR322 at 15°C. At this temperature, the oligonucleotides were in double-stranded form. The octanucleotide was cleaved at a slower rate than the other

Table I Sequences and Tm values of oligonucleotides

Sequences	Tm (°C)
AAGTACTT	14.3
CGAGTACTCG	41.0
CGCAGTACTGCG	52.3
GTCGCAGTACTGCGAC	60.9

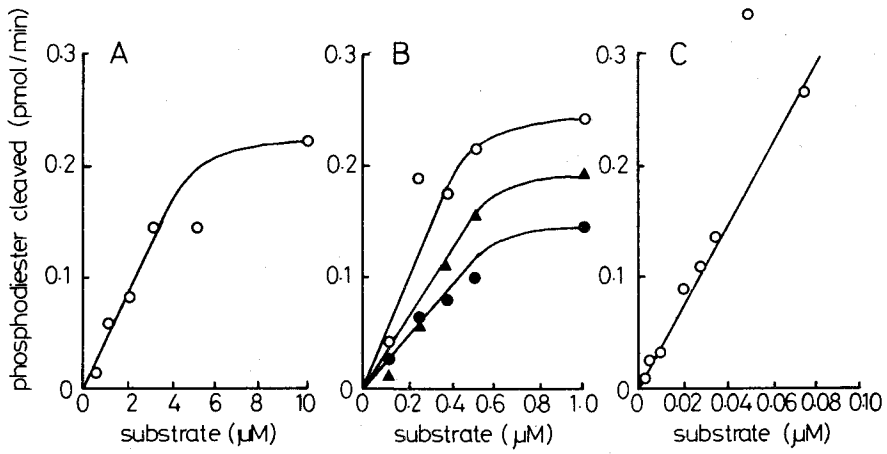

Figure 2. Effect of substrate concentration on the rate of ScaI cleavage. A; Octanucleotide was used as substrate. B; Deca- (O), dodeca- (●), and hexadecanucleotide (▲), were used as substrates. C; pBR322 was used as substrate.

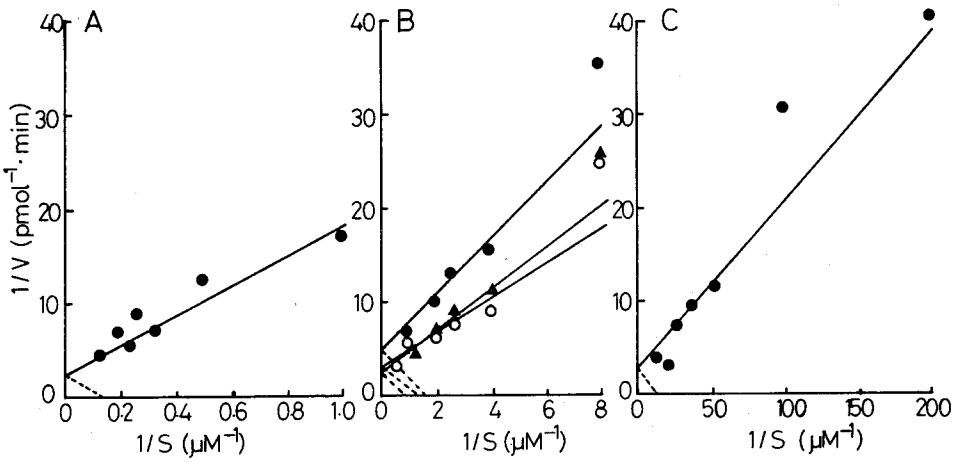

Figure 3. Double-reciprocal plots of rate of phosphodiester bond cleavage as a function of substrate concentration. A; Octa- nucleotide was used as substrate. B; Deca- (▲), dodeca- (●), and hexadecanucleotide (O), were used as substrates. C; pBR322 was used as substrate.

oligonucleotides (Fig. 2). Among the deca-, dodeca-, and hexadecanucleotides, there was little difference in the cleavage rate, and no correlation was observed between the cleavage rate and the chain length. Km and Vmax values of the cleavage by ScaI endonuclease of the oligonucleotides and pBR322 were calculated in order to explore the kinetic properties of the enzyme. ScaI followed Michaelis-Menten kinetics (Fig. 3). The apparent kinetic constants estimated from double-reciprocal plots are given in Table II. Although the Km values were different, Vmax values were very similar. Km for the octanucleotide was higher than those for oligonucleotides 10 bp or longer by one order of magnitude, and for pBR322 by two. This means that the octanucleotide was a poor substrate, and that oligonucleotides 10 bp or longer were more suitable.

Table II Kinetic parameters of ScaI endonuclease cleavage for oligonucleotides and pBR322

Substrates	Vmax(pmol/min)	Km(μM)
AAGTACTT	0.40	6.90
CGAGTACTCG	0.36	0.74
CGCAGTACTGCG	0.22	0.66
GTCGCAGTACTGCGAC	0.42	1.02
pBR322	0.37	0.066

DISCUSSION

I purified restriction endonuclease, *Sca*I, which was essentially free from the activity of other endonucleases. Using synthetic oligonucleotides as substrates, I confirmed that the enzyme recognized the sequence 5'-AGT↓ACT-3', cleaving where indicated by the arrow. Results were identical for the cleavage of large DNA substrates (32).

The octanucleotide was not cleaved efficiently. *Km* for 8 bp was 10 times higher than those for oligonucleotides 10 bp or longer. The interaction of restriction endonucleases with chemically synthesized octanucleotides has been investigated by Greene *et al*. (26) using *Eco*RI and by Dwyer-Hallquist *et al*. (36) using *Hpa*I. The *Km* of *Sca*I for AAGTACTT obtained in my experiment was of the same order as that of *Eco*RI for TGAATTCA and *Hpa*I for GGTTAACC. This is the first report comparing the affinities of a restriction endonuclease to oligonucleotides of various lengths. Recently, single-crystal X-ray diffraction studies showed that the octanucleotides GGCCGGCC (37) and GGTATACC (38) are in the *A* form and that the dodecanucleotide CGCGAATTCGCG (39,40) is in the *B* form. One possible explanation for the high *Km* for AGATATCT obtained here is that the oligonucleotide is in the *A* form, with which restriction endonucleases may not interact efficiently. I have seen that restriction endonucleases, such as *Bgl*II, *Pst*I, *Bam*HI, and their isoschizomers that recognize 6 bp, do not cleave synthetic octanucleotides but do cleave synthetic decanucleotides

(data not shown). I did not measure the affinity of the enzymes to those oligonucleotides, but I presume that the Km values of these enzymes for the octanucleotide would be higher than for the decanucleotide by one order of magnitude or more. In chapter I, I reported that the cleavage site of restriction endonucleases can be determined easily and rapidly by using synthetic oligonucleotides. The results shown here suggest that oligonucleotides which are 10 bp or longer should be used for this purpose.

SUMMARY

The kinetic constants of the site-specific endonuclease, ScaI, for various substrates were determined. I estimated Vmax and Km for octa-, deca-, dodeca-, and hexadecanucleotides and for plasmid pBR322 DNA. Vmax for these substrates were close, but Km were quite different (in decreasing order, octa- > deca-, dodeca-, hexadeca- > pBR322). The results were discussed with respect to the tertiary structure of substrate.

CHAPTER III

AccIII, a new restriction endonuclease from Acinetobacter calcoaceticus

Although many site-specific restriction endonucleases have been isolated from various kinds of bacteria, some palindromic sequences, which are not yet found to be recognized by known restriction endonucleases, still exist (7). The two restriction endonucleases, AccI and AccII, in Acinetobacter calcoaceticus have been isolated and their recognition sequences and cleavage sites have been reported (chapter I, 41). The presence of a third activity was suggested by Roberts but its details were unknown (41). However, upon reexamining A. calcoaceticus, I found a new endonuclease, AccIII, which recognized a new palindromic sequence and some of its properties were investigated.

MATERIALS AND METHODS

Bacterial strains and culture conditions A. calcoaceticus (kindly denoted M. Takanami) was used. Cells were aerobically incubated at 37°C in L-broth (1% bacto-tryptone, 0.5% yeast extract, 0.1% glucose, and 0.5% NaCl, pH 7.2) and harvested by centrifugation when it reached the late logarithmic phase. The

yield of cells per liter culture was about 8 g (wet weight).

DNA and enzymes λ phage DNA (Dam^+ λ-DNA) was prepared from *E. coli* K12 W3350 (λcI857 S7) lysogen by phenol treatment of phage particles banded in CsCl gradient according to the procedure of Thomas and Davis (28). N^6-Methyl-adenine-free λ-DNA (Dam^- λ-DNA) was purchased from New England Biolabs. Dam^+ and Dam^- pBR322 DNAs were isolated from *E. coli* C600 and GM33 (CGSC strain kindly supplied by A. Oka) cells by the procedure of Guerry *et al.*(16), respectively. ϕX174 RFI DNA was isolated from *E. coli* Cn infected with ϕX174 am3 by a modification of the method of Ueda *et al.* (42). Decanucleotide d(GTTCCGGAAC) was synthesized by the solid phase method (19). Adenovirus-2 and SV40 DNA were purchased from Bethesda Research Laboratories, Inc. Restriction endonucleases *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I, and *Hin*cII were prepared from *E. coli* RY13 (43), *Haemophilus influenzae* Rd (44), *E. coli* ED8654 carrying pBR322 with a *Pst*I gene insertion (45), *Streptomyces albus* G (46), and *Haemophilus influenzae* Rc (47), respectively. *Acc*I and *Acc*II were products of Takara Shuzo Co., Ltd.

Assay of *Acc*III endonuclease Enzyme activity was measured in a reaction mixture (50 μl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 60 mM NaCl, 7 mM 2-mercaptoethanol, 0.01% BSA, and 1 μg of DNA. One unit was defined as the amount of enzyme required to digest 1 μg of λ-DNA completely in 60 min at 37°C.

Purification of *Acc*III 349 g of cells were suspended in 10 mM potassium phosphate, pH 7.5, containing 10 mM 2-mercaptoethanol,

and disrupted by sonication. Debris was removed by centrifugation (10^5 x g for 1 h). The supernatant was treated with 50% (w/v) ammonium sulfate and the precipitate was collected and suspended in KP buffer (10 mM potassium phosphate, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerol). After dialysis against this buffer, the enzyme solution was applied on to a phosphocellulose (Whatman P11) column (bed volume 300 ml), and eluted with a linear gradient of 0-1.0 M KCl in KP buffer. The AccIII was eluted at 0.65-0.7 M KCl, and was separated from the other two endonucleases, AccI and AccII. The AccIII fractions were pooled, dialyzed, and applied on to an Affi-Gel Blue agarose (Bio-Rad) column (bed volume 4 ml). The AccIII fraction was passed through the column. The active fractions were applied on to a DEAE-cellulose (Whatman DE52) column (bed volume 10 ml), and eluted with a linear gradient of 0-1.0 M KCl in KP buffer. The active fractions (0.20-0.25 M KCl) were pooled, dialyzed, applied on to a heparin-Sepharose (Pharmacia CL-6B) column (bed volume 4 ml), washed with 0.7 M KCl in KP buffer and eluted with 1.0 M KCl in KP buffer. The active fractions were pooled, dialyzed, and finally applied on to an Aminohexyl agarose (BRL) column (bed volume 4 ml), and eluted with a linear gradient of 0-1.5 M KCl in KP buffer. The active fractions (0.32-0.65 M KCl) were pooled, and concentrated with polyethyleneglycol. An equal volume of glycerol was added and the enzyme preparation was stored at -20°C.

Determination of cleavage site for AccIII The cleavage site of AccIII was determined using the synthetic decanucleotide

d(GTTCCGGAAC) by the method described in chapter I. AccIII (0.4 units) was added to a reaction mixture (20 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 60 mM NaCl, and 0.45 µM 5'-^{32}P-labeled oligonucleotides, and the solution was incubated at 37°C for 30 to 120 min. The resulting oligonucleotides were then separated by homochromatography (20) on a DEAE-cellulose thin layer plate (Masherey-Nagel CEL 300 DEAE/HR-2/15) using Homomixture III and detected by autoradiography.

RESULTS

Optimal conditions for AccIII activity Effects of salt concentration and temperature on the activity of AccIII were examined and were compared with those of AccI and II that were produced in the same organism. As shown in Fig. 1, optimum conditions for AccIII were different from those for AccI and II. Maximum AccIII activity was obtained around 60-65°C in a buffer containing 150 mM KCl or NaCl, 20 mM $MgCl_2$, at pH 8.5.

Determination of AccIII recognition sequence Phage and plasmid DNA of known sequences were incubated with AccIII. Both Dam$^+$ and Dam$^-$ λ-DNA were cleaved with AccIII at 20 or more sites, but the patterns were different from each other at several fragments (Fig. 2). AccIII did not cleave SV40, ϕX174 or Dam$^+$ pBR322, but cleaved Dam$^-$ pBR322 and Adenovirus-2 at a unique site and eight sites, respectively (Fig. 3). As mentioned for

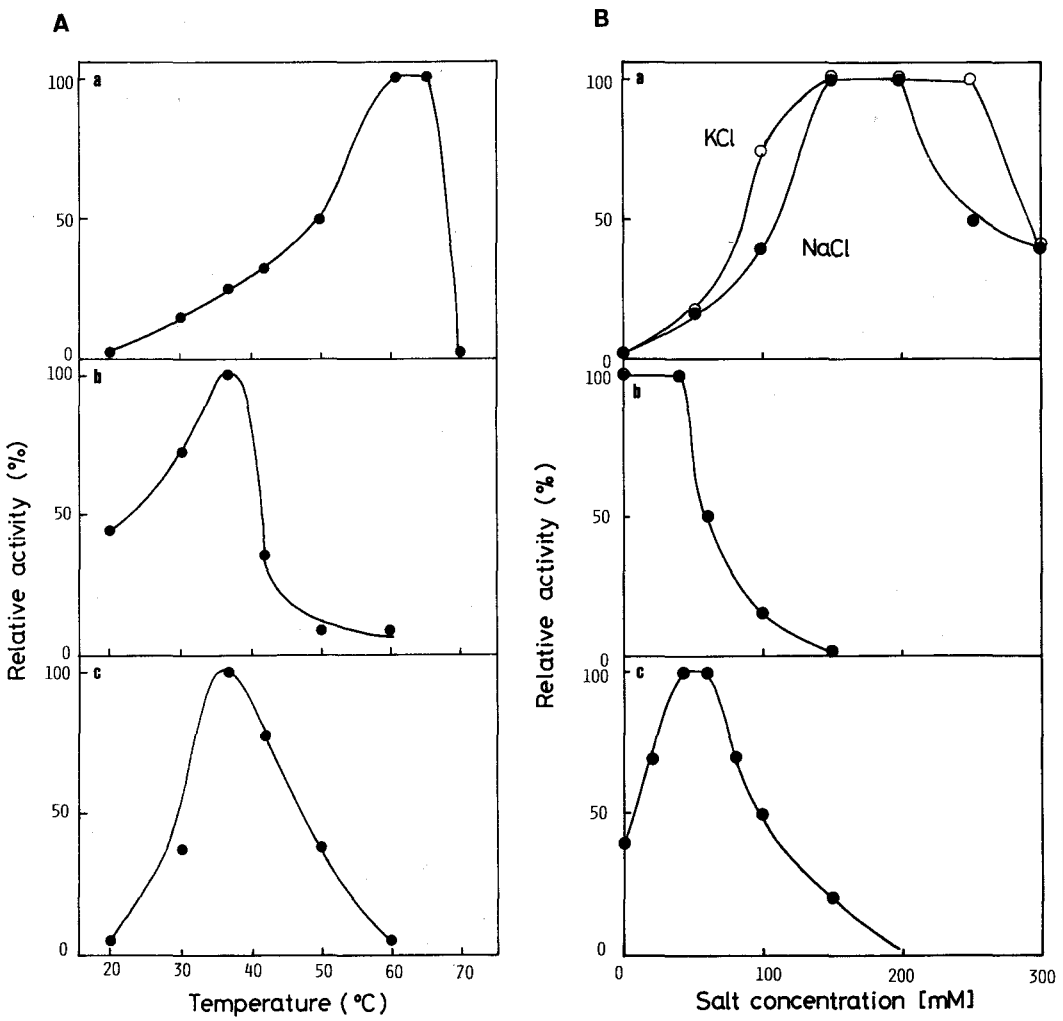

Figure 1. Effects of temperature and salt on the activity of AccI, II, and III. A, Effect of temperature. Enzyme activity was measured in the reaction mixture (50 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM 2-mercaptoethanol, 7 mM $MgCl_2$, and 0.01% BSA containing 150 mM NaCl (AccIII(a)) or 60 mM NaCl (AccII(b) and AccI(c)), respectively. B, Effect of salt. Enzyme activity was measured in the reaction mixture (50 µl) containing 10 mM Tris-HCl, pH 7.5, 7 mM 2-mercaptoethanol, 7 mM $MgCl_2$, and 0.01% BSA. Assay temperature was 60°C (AccIII(a)) or 37°C (AccII(b) and AccI(c)), respectively.

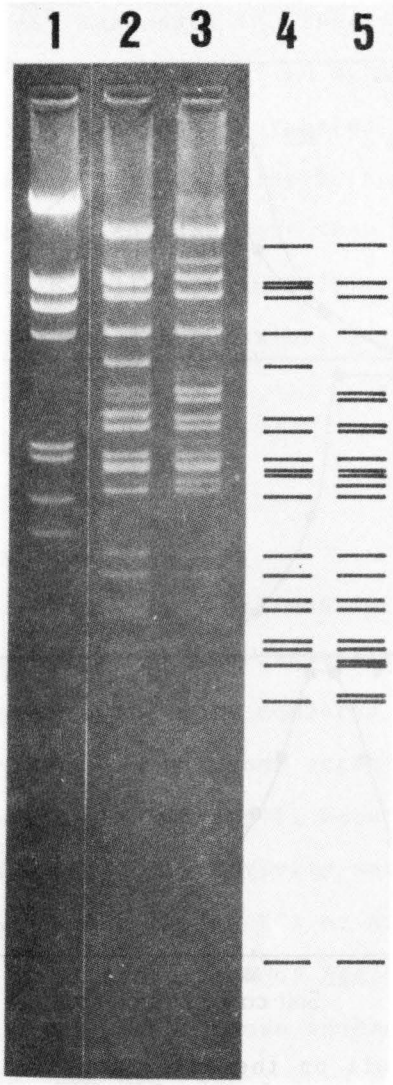

Figure 2. Agarose gel electrophoresis of Dam^+ and Dam^- λ-DNA cleaved with AccIII. Lane 1 is molecular weight marker (λ-HindIII,EcoRI fragments). AccIII was incubated with Dam^+ (lane 2) and Dam^- (lane 3) λ-DNA. Lane 4 (Dam^+) and lane 5 (Dam^-) are computed patterns of AccIII cleaved λ-DNA.

MflI (48), when the dam modification sequence, GATC (49), completely overlapped with the MflI recognition sequence, Dam^+ λ-DNA was not cleaved at all. However, when the dam modification sequence partly overlapped, Dam^+ λ-DNA was cleaved but the patterns were different from those of Dam^- λ-DNA, e.g. ClaI (50).

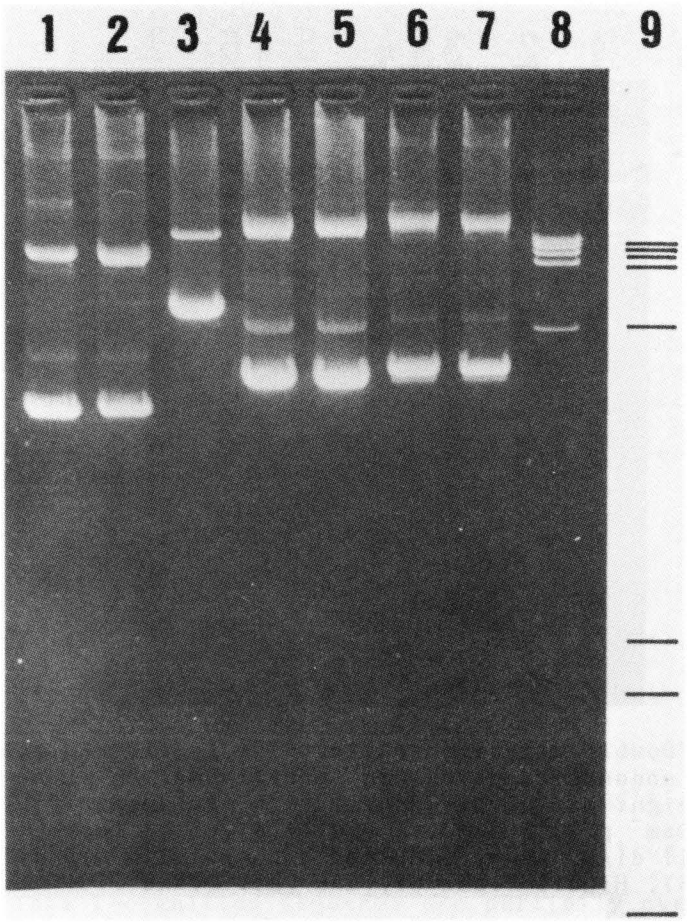

Figure 3. Agarose gel electrophoresis of DNA cleaved with AccIII. Various DNA were incubated with AccIII. Lane 1, Dam^+ pBR322; lane 2, Dam^+ pBR322+AccIII; lane 3, Dam^- pBR322+AccIII; lane 4, SV40; lane 5, SV40+AccIII; lane 6, øX174; lane 7 øX174+AccIII; lane 8, Adenovirus-2+AccIII (The small molecular bands of lane 8 were too small to be visible.); lane 9, computed pattern of lane 8.

I presumed that GATC partly overlapped at some of the AccIII recognition sites. Double digestion of Dam^- pBR322 with AccIII and with each of HindIII, SalI, HincII, and PstI, showed that a unique cleavage site was assigned around 1700 bp (Fig. 4).

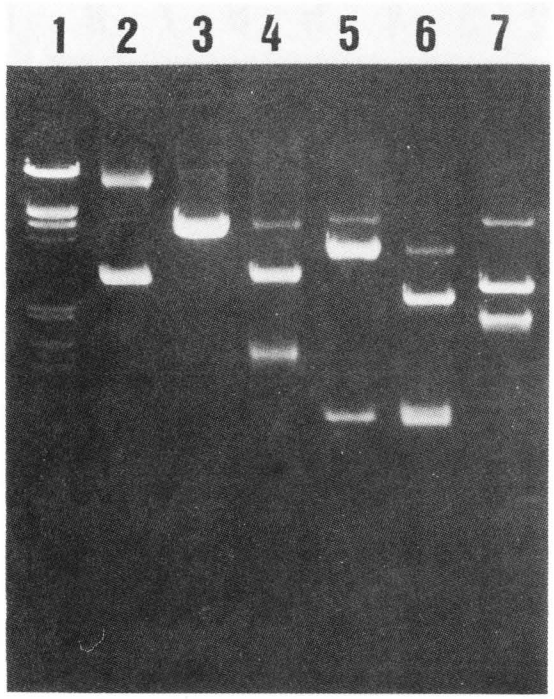

Figure 4. Double digestion patterns of AccIII and other restriction endonucleases on Dam⁻ pBR322 DNA. Lane 1 is molecular weight marker (λ-HindIII,EcoRI fragments). Lane 2 and lane 3 are Dam⁻ pBR322 and its AccIII digests, respectively. Dam⁻ pBR322 AccIII digests were further cleaved with HindIII (lane 4), SalI (lane 5), HincII (lane 6), and PstI (lane 7).

Furthermore, the computer search of the GATC sequence around the 1700 bp region indicates that GATC appeared three times in the 1459 to 3040 bp region (51,52). However, since the TCCGGATC which is present at 1664 to 1671 was closest to the unique cleavage site which was assigned around 1700 bp as mentioned above, it is suggested that AccIII recognizes the palindromic sequence TCCGGA, which is present at the 1664 to 1669 bp region.

A computer search of TCCGGA through the sequence of λ,

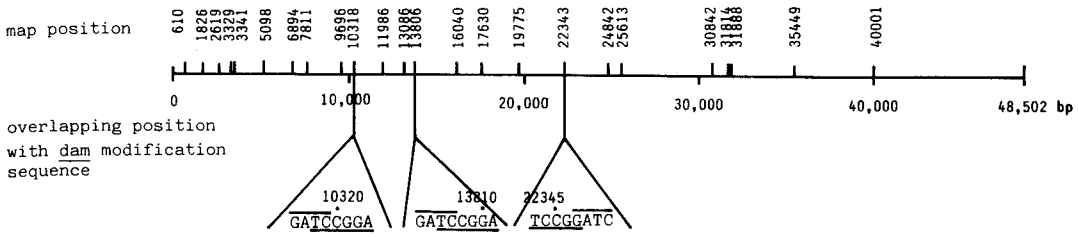

Figure 5. Mapping of AccIII restriction sites in λ-DNA. Numbering of the nucleotide sequence begins with the first base of the left end. The map positions are shown as the number of the first 5' base in the AccIII recognition sequence. The positions partly overlapped with dam methylase modification sequence are indicated below the line.

pBR322, SV40, øX174, and Adenovirus-2 indicates that this sequence should respectively occur at 24 (51,53), 1 (52), 0 (54), 0 (55), and 8 (58) different sites. These are compatible with the digestion patterns of these DNA with AccIII, as shown in Fig. 2 and 3. The recognition sequence was confirmed on a complete AccIII map of λ-DNA (53). As shown in Fig. 5, the AccIII recognition sequence and the dam methylase recognition sequence are partially overlapped at three positions, 10320, 13810, and 22345 bp. Knowing that AccIII activity was inhibited at the three positions, the fragment size of both Dam^+ and Dam^- λ-DNA were computed and are shown in Fig. 2 (lane 4 and 5). As expected, the patterns produced by AccIII cleavage were the same as those computed from the nucleotide sequence.

Determination of AccIII cleavage site The cleavage site of AccIII was identified using a synthetic oligonucleotide d(GTTCCGGAAC). The oligonucleotide was labeled with ^{32}P at the

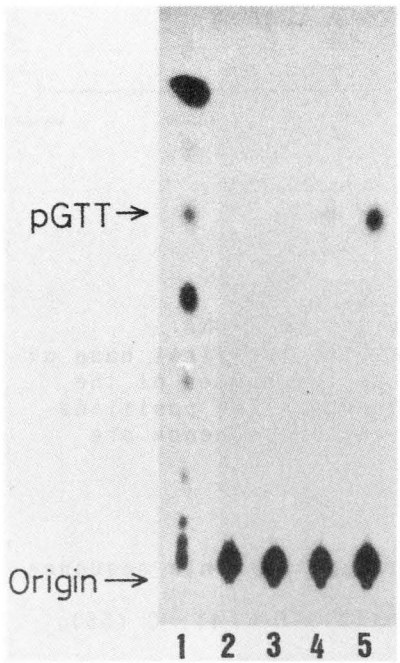

Figure 6. Identification of oligonucleotides produced from synthetic decanucleotide by digestion with AccIII. Lane 1 is of authentic markers obtained by digestion with venom phosphodiesterase. The decanucleotide was digested with AccIIII for different times (lane 2, 0 min; lane 3, 30 min; lane 4, 60 min; lane 5, 120 min).

5'-terminus, and then digested with AccIII. The resulting labeled oligonucleotides were then separated and detected by autoradiography. The trinucleotide ^{32}P-GTT was found to be the product of AccIII digestion (Fig. 6). This shows that AccIII cleaves the substrate at GTT↓CCGGAAC, where indicated by the arrow.

DISCUSSION

This is the first report of a type II restriction endonuclease whose recognition sequence is TCCGGA. AccIII required different conditions for maximum activity from those of

other two endonucleases in A. calcoaceticus. The optimum salt concentration for AccIII activity was 150 mM, at which concentration AccI and II activities were considerably inhibited. Although the optimum temperature for AccI and II's maximum activity and the growth of A. calcoaceticus was around 30-37°C (57). AccIII showed maximum activity around 60-65°C. These differences show that in A. calcoaceticus, AccIII may play a distinct role from those of the other two enzymes.

Inhibition of cleavage due to overlapping E. coli dam or dcm modification sites have been observed for several restriction endonucleases (7). Partly overlapping with dam modification resulted in resistance to cleavage by ClaI (GATCGAT), HphI (GGTGATC), MboII (GAAGATC), NruI (GATCGCGA), TaqI (GATCGA), and XbaI (GATCTAGA) (9). Here, AccIII was found to be inhibited by such a modification.

AccIII cleaves between T and C in the recognition sequence, TCCGGA, and produces 5'-protruding tetranucleotides, CCGG. The AccIII digestion products could be ligated directly with Cfr10I ((A/G)↓CCGG(T/C)), Cfr9I, XcyI, and XmaI (C↓CCGGG) fragments, which all have the CCGG sequence at the 5'-termini (7).

SUMMARY

A new site-specific restriction endonuclease, AccIII, was isolated from A. calcoaceticus. AccIII recognizes

T↓CCGGA and cleaves at the position shown by the arrow. AccIII activity was inhibited by adenine methylation at the overlapping dam methylase recognition sequence.

CHAPTER IV

Organization and nucleotide sequence of the restriction and modification genes of Flavobacterium okeanokoites

The genes of variety of type II restriction-modification systems have already been cloned, and their organization and nucleotide sequences have been analysed (58-65). Some of the gene products have been purified, and for the EcoRI endonuclease, the protein conformation that specifically interacts with the recognition sequence has been elucidated (16). However, all of these enzymes analysed belong to the type II systems that recognize symmetrical sequences. So far, there is no report on the gene structure of restriction-modification systems that recognize asymmetric sequences. The only paper available is on the cloning of the methylase genes of the FokI and HgaI systems (67). I am interested in elucidating the mechanisms of recognition of asymmetric sequences and of cleavages away from the recognition sequences. Experiments were designed to clone the systems that belong to this category, and a DNA fragment carrying the genes of the complete FokI system was successfully isolated from a cosmid library of Flavobacterium okeanokoites.

RFokI recognizes asymmetric pentanucleotides in double-stranded DNA, 5'-GGATG-3' in one strand and 3'-CCTAC-5' in the other, and introduces staggered cleavages at sites downstream from

the recognition site (68). MFokI modifies DNA, and this renders the DNA resistant to digestion by RFokI. I deduced the nucleotide sequence of the cloned fragment, and assigned the coding regions for MFokI and RFokI on the basis of deletion analysis. The cloned genes were expressed in E. coli cells, from which both enzymes could be purified to homogeneity.

MATERIALS AND METHODS

Strains F. okeanokoites IFO 12536 was obtained from the Institute for Fermentation, Osaka, Japan. E. coli strains HB101 (69) and JM109 (70), and plasmids pHC79 (71) and pUC18 (70) were the host-vector systems used for cloning. M13mp18 and mp19 were used for DNA sequencing (70).

Enzymes and chemicals Restriction endonucleases, in vitro packaging kits, M13 sequencing kits and Kilo-sequence deletion kits were products of Takara Shuzo Co., Ltd., and used according to the manufacturer's instructions. [α-^{32}P]dCTP (400 mCi/mmol) was purchased from Amersham Inc.

Construction of cosmid libraries F. okeanokoites DNA was purified by the procedure of Thomas et al. (72). Partial Sau3AI digests of F. okeanokoites DNA (100 μg) were fractionated by centrifugation on a 5-25% sucrose gradient, and fractions predominantly containing the fragments of 35-50 kb were collected. The sized Sau3AI-fragments (1.7 μg) were ligated to a phosphatase-

treated BamHI digest of pHC79 (0.3 µg), and packaged into λ phage capsids with the in vitro packaging kit. After adsorption, cells were plated onto L-agar plates containing 50 µg/ml ampicillin. Transductants were scraped together, inoculated into 200 ml of L-broth containing 100 µg/ml ampicillin, and grown to saturation, from which cosmid DNA was prepared by the alkaline method (73).

Selection of FokI modification clones FokI modification clones were selected by the procedure suggested by Mann et al. (11). Purified cosmid DNA (10 µg) was digested for 1 h at 37°C with 100 units of RFokI and then dephosphorylated. The digest was introduced into competent cells (E. coli HB101), and cells were plated onto L-agar plates containing 50 µg/ml ampicillin. Plasmids harbored in individual transformants were purified by the alkaline miniprep-procedure (74), and subjected to restriction analysis.

Subcloning of cosmid clones Cosmid clones were partially digested with HindIII, and ligated to phosphatase-treated HindIII digests of pUC18. The products were introduced into E. coli JM109 cells and plated onto L-agar plates containing 50 µg/ml ampicillin. The transformants were scraped together and inoculated into 100 ml of L-broth containing 100 µg/ml ampicillin. Cells were grown to saturation, from which plasmid DNA was prepared by the alkaline method. The DNA was digested with RFokI, and again introduced into E. coli JM109. The cells were plated onto L-agar plates containing 50 µg/ml ampicillin and lac indicators.

Assay of enzyme activities The R*Fok*I activity *in vivo* was estimated from the plating efficiency of λ phage. The R*Fok*I activity *in vitro* was assayed by incubation for appropriate periods at 37°C in 25 µl of reaction mixtures containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 0.5 µg of λ-DNA. For assay of the M*Fok*I activity *in vitro*, the sample to test was incubated for 1 h at 37°C in 20 µl of a reaction mixture containing 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM 2-mercaptoethanol, 0.1 mM AdoMet, and 1 µg of λ-DNA. The reaction was terminated by heating for 10 min at 70°C. Following the addition of 16.5 µl of 50 mM Tris-HCl, pH 7.5, 40 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 132 mM NaCl, and 6 units of R*Fok*I, the solution was incubated for 1 h at 37°C, and aliquots were analysed by gel electrophoresis. Modification was judged by the lack of digestion. The M*Fok*I activity *in vivo* was assayed by susceptibility against R*Fok*I of plasmid DNA which was prepared from cells to test.

Deletion analysis Deletions in plasmid pFokMR5.2a and pFokMR5.2b (Fig. 1) were introduced by use of the Kilo-sequence deletion kit. Plasmid DNA (about 10 µg) was digested at the *Kpn*I and *Sal*I sites in the multicloning sites. The digest was dissolved in 100 µl of ExoIII buffer and treated with 180 units of exonuclease III at 37°C. Samples (10 µl) were removed at 1 min intervals and added to a tube containing 100 µl of mung bean nuclease buffer. After incubation for 5 min at 65°C, the DNA was treated with 50 units of mung bean nuclease for 1 h at 37°C. The

termini of DNA were converted to blunt ends, and after circularization with ligase, the DNA was used for transformation of JM109. Deletion derivatives of plasmid pFokM2.4 were constructed in the same way, except that cleavages were introduced at the SacI site in the multicloning sites and the ClaI site in the insert (Fig. 1 and 3).

Purification of enzymes About 150 g of cells harboring pFokMR5.2a were suspended in 400 ml of 10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 10 mM 2-mercaptoethanol, and 500 μg/ml lysozyme, left on ice for 1 h, and then sonicated. The cell debris was removed by centrifugation (10^5 x g for 1 h). To the supernatant was added NaCl to 0.1 M, and then a 10% stock solution of polyethyleneimine to 1%. After removal of the precipitates by centrifugation, solid ammonium sulfate was added to 70% saturation. The resulting precipitates were collected by centrifugation, dissolved in 10 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerol (buffer A) containing 0.15 M KCl, and dialyzed against the same buffer. The solution was applied onto a phosphocellulose (Whatman P11) column (5 cm x 7.5 cm), and chromatographed with a linear KCl gradient in buffer A (0.15 M to 1 M, 1500 ml). The RFokI and MFokI activities were eluted at 0.25-0.3 M KCl and 0.4-0.7 M KCl, respectively. The methylase fraction from the phosphocellulose column was dialyzed against buffer A, put on a DEAE-cellulose (Whatman DE52) column (2 cm x 10 cm), and eluted with a linear gradient of KCl in buffer A (0 to 0.5 M, 200 ml). The active fraction, which was eluted at 0.16-0.38 M KCl, was then

put on a hydroxylapatite (Clarkson) column (1 cm x 6 cm), and eluted with a linear gradient of potassium phosphate buffer, pH 7.5 (0.01 M to 0.5 M, 200 ml). The activity was eluted at the buffer concentration of 0.04 M to 0.11 M. The enzyme fractions were pooled, concentrated by Centriflo (Amicon), and stored at 4°C. The endonuclease fraction from the phosphocellulose column was dialyzed against buffer A, and applied to a DEAE-cellulose column (1.5 cm x 6 cm), equilibrated with buffer A. The activity was recovered in the flow-through fraction. The fraction was adsorbed to an Affi-Gel Blue agarose (Bio-Rad) column (0.7 cm x 25 cm), and eluted with a linear gradient of KCl in buffer A (0 to 1 M, 200 ml). The activity peak, which was eluted at 0.1-0.35 M KCl, was then put on a hydroxylapatite column (0.15 cm x 6 cm), and eluted with a linear gradient of potassium phosphate buffer, pH 7.5 (0.1 to 0.5 M, 200 ml). The active fraction, eluted at the buffer concentration of 0.15-0.21 M, was further fractionated on a Sephadex G-100 (Pharmacia) column as in the next section, concentrated by Centriflo, and stored at 4°C. There was little loss of activities of both enzymes during storage for 3 months.

Molecular weight measurement by gel electrophoresis

Approximate molecular weights in the native form of enzymes were estimated by gel filtration through a Sephadex G-100 column (2 cm x 70 cm). The activity peak of each enzyme in the final purification step and a mixture of the calibration standards (catalase, BSA, ovalbumin, chymotrypsinogen A, cytochrome C; 0.5 ml each) were applied onto the columns, and gel filtration was

carried out with buffer A containing 0.15 M KCl. Fractions (2 ml) were collected, and aliquots were used for analysis of the protein concentration and enzyme activities.

Analysis of N-terminal amino acid sequences The methylase in the final purification step was blotted from SDS-polyacrylamide gel to Polyvinylidene difluoride membranes (Millipore) by the method of Matsudaira (75). The sequencial degradation of protein was performed with an Applied Biosystems Model 470A, equipped with an on-line HPLC apparatus Model 120A.

DNA sequencing The dideoxynucleotide chain-termination method (76,77) was used for detemination of the entire sequence of the 4.3 kb region, which encompassed both of the FokI genes. The subfragments generated were cloned into appropriate sites of M13mp18 or M13mp19 to provide templates, and the sequences were deduced from the data for both strands.

RESULTS

Isolation of FokI restriction-modification genes First, experiments were designed to isolate recombinants carrying the FokI modification gene by the survivor-selection method from a cosmid library that had been intensively digested with RFokI. About 1,400 colonies carrying plasmids were surveyed, and finally five colonies randomly picked up were found to carry recombinants resistant to RFokI digestion. These clones were named pFokcos13,

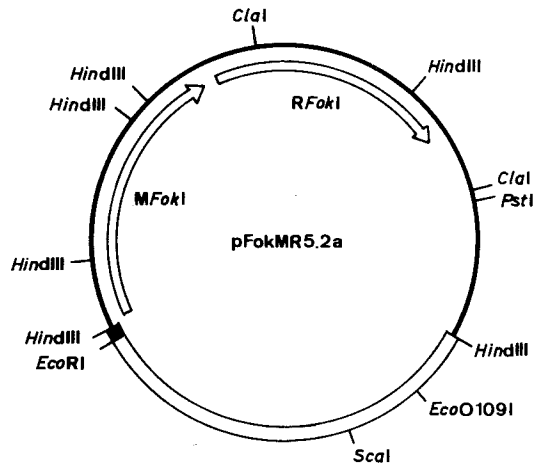

Figure 1. Restriction maps of plasmid pFokMR5.2a encoding the FokI restriction-modification genes. Open and closed boxes and the line represent the pUC18 moiety, its multicloning sites, and the 5.2 kb insert, respectively. The areas and direction of RFokI and MFokI genes are shown by open arrows.

14, 15, 16, and 25. When the crude extracts from cells carrying these clones were assayed for the MFokI activity, activity was detected in all the extracts. Then, the RFokI activity was analysed by both in vitro and in vivo assay procedures. Plating efficiencies of λ phage for cells carrying respective recombinants were reduced to about half of the efficiency for control cells carrying no cosmid. Some activity of the specific endonuclease was also detected in all of the crude extracts. Among these clones, pFokcos16, containing the 38 kb insert, expressed the highest methylase activity. Thus, subclones carrying the MFokI gene were screened from its partial HindIII digests, and five independent colonies that carried plasmids resistant to RFokI digestion were isolated. Endonuclease and methylase activities in

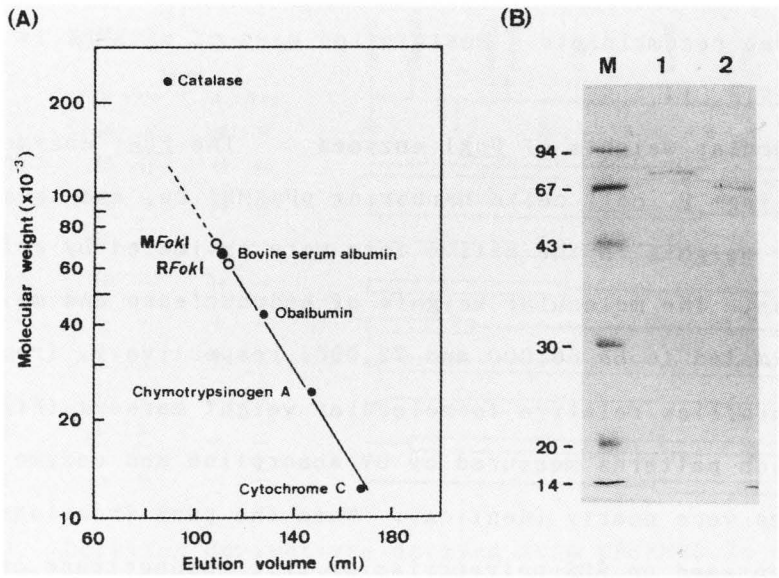

Figure 2. Estimation of molecular weights of FokI enzymes by gel filtration (A) and SDS-polyacrylamide gel electrophoresis (B). (A) Activity peaks in the final purification step of RFokI and MFokI and molecular-weight markers were chromatographed on Sephadex G-100 columns, and elution volumes at peak positions of the different proteins were plotted as the logarithmic function of molecular weights. (B) Molecular-weight markers (lane M) and activity peaks of MFokI (lane 1) and RFokI (lane 2), eluted from the Sephadex G-100 columns, were electrophoresed on gradient SDS-polyacrylamide gels (10% to 20%), and stained with Coomassie brilliant blue. Numbers by the side of lane M are molecular masses in kDa of marker proteins.

cells from these colonies were examined by both *in vivo* and *in vitro* assay procedures. Among the colonies examined, those carrying plasmids, named pFokMR5.2a and pFokMR5.2b, expressed both methylase and endonuclease activities to the same significant extent. When the inserts were analysed, these two plasmids carried the same 5.2 kb fragment but in reverse orientations. The plating efficiency of λ phage for cells carrying either pFokMR5.2a

or pFokMR5.2b was reduced to 10^{-2} of that for control cells carrying no recombinants. Restriction maps of pFokMR5.2a are shown in Fig. 1.

Molecular weights of FokI enzymes The FokI enzymes were purified from E. coli cells harboring pFokMR5.2a, and their molecular weights in the native form were estimated by gel filtration. The molecular weights of endonuclease and methylase were estimated to be 66,000 and 72,000, respectively, from the elution profiles relative to molecular weight markers (Fig. 2A). The elution patterns measured by UV absorption and enzyme activities were nearly identical. When the peak fractions were electrophoresed on SDS-polyacrylamide gel, endonuclease and methylase gave single bands of 66 and 75 kDa (Fig. 2B, lanes 1 and 2), in accordance with the estimates by gel filtration. The results also indicated that the monomeric forms of both enzymes were catalytically active.

Gene organization and nucleotide sequences To analyse the organization of the two genes encoded in the 5.2 kb fragment, series of deletion mutants were constructed from either pFokMR5.2a or pFokMR5.2b, and their abilities to confer restriction and modification phenotypes on E. coli cells were examined. The results are summarized in Fig. 3. pFokMR3.9 was the smallest one that conferred both restriction-modification phenotypes, and pFokl.9 conferred neither the restriction nor the modification phenotype. pFokM3.7 and pFokM2.4 could confer the modification phenotype at a level equivalent to that of pFokMR5.2, but not any

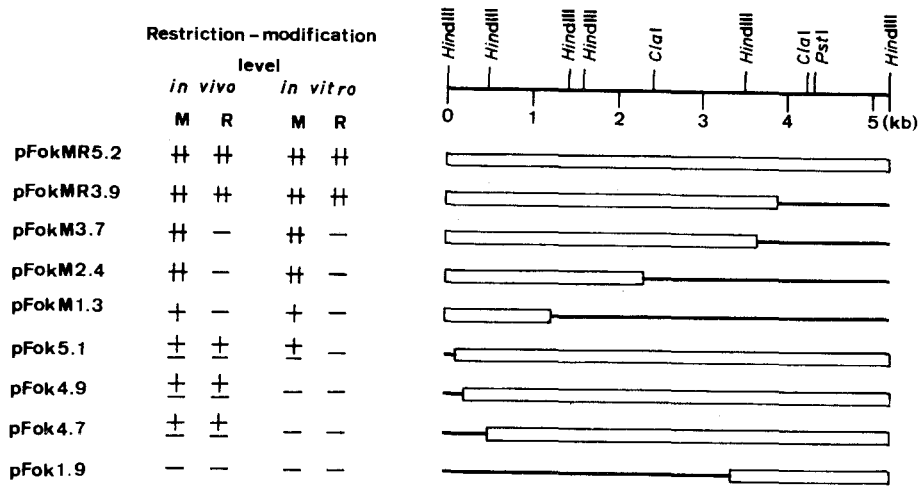

Figure 3. Deletion derivatives derived from pFokMR5.2a and pFokMR5.2b. The horizontal coordinate in the right part of the figure represents the physical map of the 5.2 kb insert. Below the map, deleted and remaining portions of the insert in derivatives are indicated by lines and open boxes, respectively. In the left part, the name of the derivatives and RFokI (R) and MFokI (M) activities, which were expressed in cells, are indicated. The activities were assayed both in vivo and in vitro, and represented by the approximate levels relative to those of pFokMR5.2. ++ , Equivalent levels; + , levels about half to one-tenth; ± , levels less than one tenth; - , no detectable activity.

restriction phenotype. pFokM1.3 also gave the modification phenotype alone, but at a decreased level. By introduction of small deletions from the leftmost end of the 5.2 kb insert (pFok5.1, pFok4.9, and pFok4.7), both restriction-modification phenotypes were identified in the in vivo assay, but to greatly reduced extents. Although little or no activity was detected in the in vitro assay, this was due to the insensitivity of the assay. Taken together with the observation that two plasmids, pFokMR5.2a and pFokMR5.2b, carrying the 5.2 kb insert in reverse

orientations expressed the same level of restriction and modification activities, it was strongly suggested that the leftmost end region was involved in regulation of gene expression. The appearance of a low level of restriction-modification phenotypes in the clones, in which the left end region was deleted, was probably due to read-through transcripts from the vector moiety.

The results of deletion analysis outlined the location of genes for the FokI restriction-modification system. On the physical map shown in Fig. 3, the regulatory region of gene expression is located near the leftmost end, followed by methylase and endonuclease genes in that order.

The DNA sequence of the 4.3 kb region in pFokMR5.2a that covers the entire MFokI and RFokI genes (the left end to the PstI site of the 5.2 kb insert) is shown in Fig. 4. Analysis of open reading frames indicated that only two could code for proteins of the expected molecular weights (Fig. 5). The location and orientation of these two open reading frames were in agreement with the results of deletion analysis. In the open reading frame assigned for the methylase gene, the ATG codon appeared at nucleotide positions 62 and 149 and the termination codon at nucleotide position 2,003. I assigned the initiation codon to the one at nucleotide position 62, because the predicted molecular weight (75,622) agreed well with the value (75,000) of MFokI estimated by SDS-PAGE (Fig. 2). In addition, an appropriate ribosome-binding sequence, AGGA, appeared 9 bp upstream of the ATG

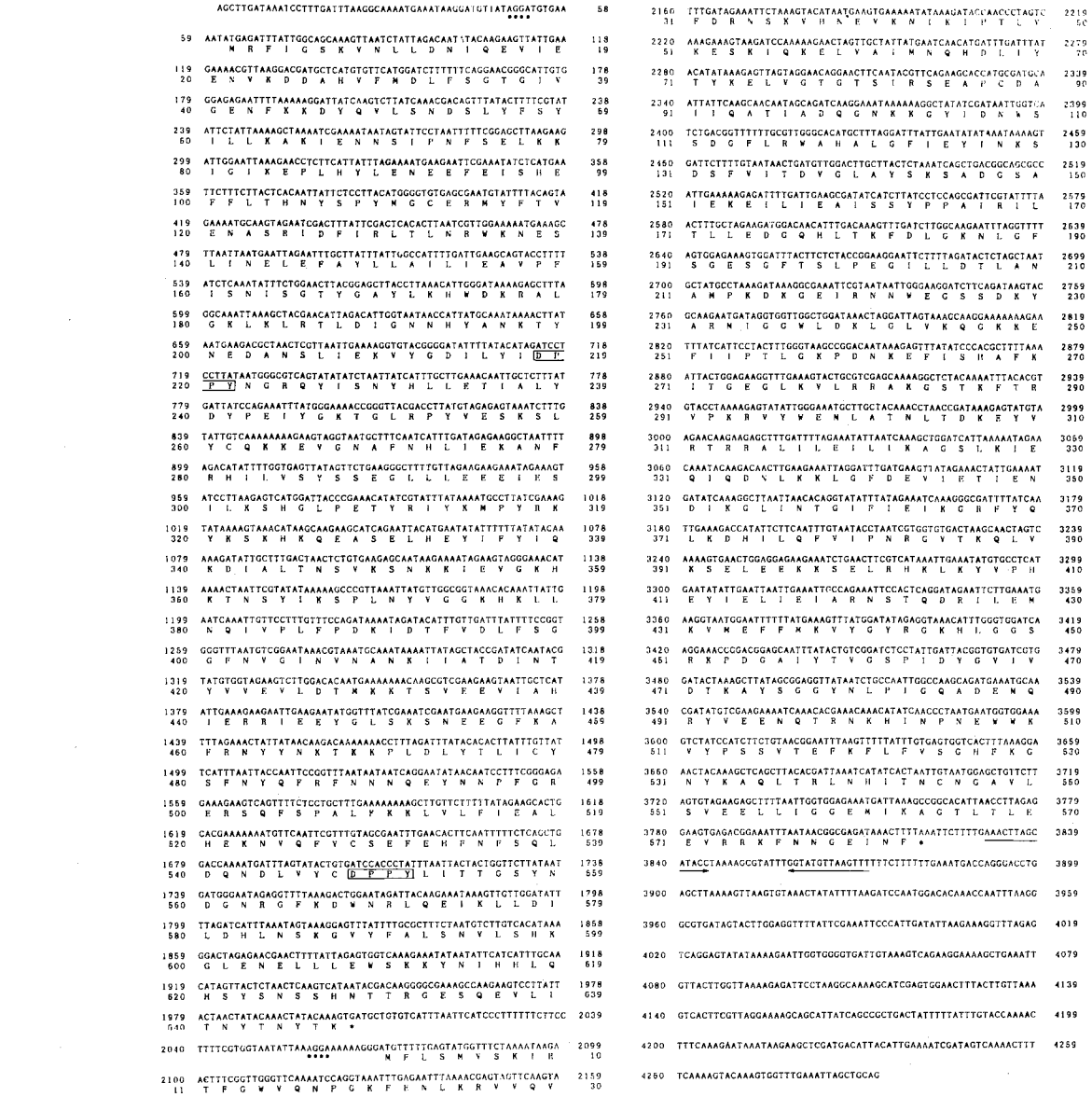

Figure 4. Entire nucleotide sequence of the 4,288 bp region from the left end to the PstI site of the 5.2 kb insert in pFokMR5.2a. The amino acid sequences assigned for RFokI and MFokI are given below the nucleotide sequence. The nucleotide sequence is numbered from the leftmost end, and the amino acid sequences from the initiation codons of each gene. The potential ribosome-binding sequences are dotted, and the adenine methylase-specific segments in MFokI are boxed. A pair of arrows show palindromic sequences characteristic of the termination signal.

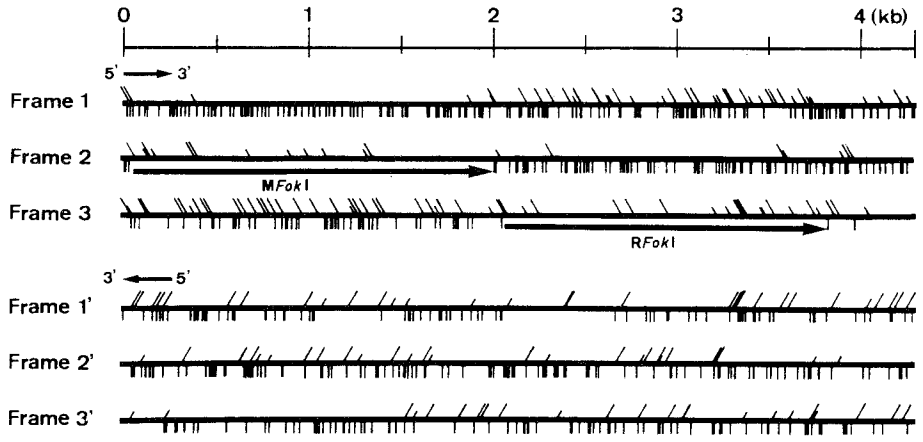

Figure 5. Codon analysis for protein synthesis. Initiation and termination codons were plotted on each of three possible frames. Long- and short-angled lines represent ATG and GTG, respectively, and vertical lines represent TAA, TAG, and TGA. The coding regions assigned for RFokI and MFokI are indicated by arrows.

codon. The start of the methylase gene was further confirmed by N-terminal amino acid analysis. The first nine amino acids of the enzyme obtained by Edman degradation were Met-Arg-Phe-Ile-Gly-Ser-Lys-Val-Asn. This sequence exactly corresponded to that predicted from the nucleotide sequence at position 62.

In the open reading frame for the RFokI gene, the ATG codon occurred at 2,070 and 2,082, and the termination codon at nucleotide position 3,819. The predicted molecular weights of protein beginning with ATG at nucleotide positions 2,070 and 2,082 were 66,216 and 65,737, respectively. Both values were close enough to the molecular weight (66,000) estimated by SDS-PAGE. I assigned the ATG codon at 2,070 to the initiation site for translation of the endonuclease gene, for only this one was accompanied by putative ribosomal-binding sequence, AGGA, in the

appropriate position. To confirm this assignment, I attempted to analyse the N-terminal amino acid sequence, but have not been successful yet.

DISCUSSION

I isolated a clone carrying the complete FokI restriction and modification genes from a cosmid library of F. okeanokoites chromosomal DNA by the selection procedure suggested by Mann et al. (11). Although cloning and restriction maps of FokI methylase have already been reported by Nwankwo and Wilson (67), the results reported here indicate that their 3.8 kb fragment does not encompass the complete methylase gene. The success in cloning of the complete genes is probably owing to the use of a cosmid that permitted packaging in vitro of large DNA fragments.

Two classes of gene organization have been observed for the type II restriction-modification systems; one is in the same orientation (EcoRI (58,59), HhaII (60), BsuRI (63), PaeR7I (64), TaqI (65), DdeI (66)), and the other in the opposite orientation (EcoRV (61), PstI (62)). The systems that belong to the former class are further divided into two groups by the order of the restriction and modification genes. The FokI system reported here, belongs to the group in the same orientation, with the modification gene preceeded.

The two recombinants, pFokMR5.2a and pFokMR5.2b, which

carried the 5.2 kb insert in reverse orientations, both expressed about the same level of modification and restriction activities. This implies that a promoter, that could be recognized by E. coli RNA polymerase, was located within the 62 bp region from the leftmost end of the insert to the first initiation codon. By computer-aided search, I tentatively assigned 5'-TTGATA-3' beginning at nucleotide position 4 and 5'-CAAAT-3' beginning at nucleotide position 28 for the -35 and -10 regions of the promoter, but only the former seemed to adhere to the -35 consensus sequence of E. coli promoters. Immediately downstream of the RFokI gene, a palindromic sequence followed by a stretch of thymine-residues is seen (Fig. 4). This could be the termination site of transcription. Although no information is available about the structure of the promoter in F. okeanokoites, it is generally accepted that prokaryotic promoters have some structural similarity. As both the restriction and modification activities were concomitantly reduced by introduction of deletions upstream of the MFokI gene, it is likely that the RFokI and MFokI genes form an operon the expression of which is regulated by a promoter upstream of the MFokI gene.

Sequence comparison of the two FokI enzymes showed little similarity except for the sequence Ile-Leu-Ile-Glu-Ala that appeared at amino acid-positions 152-156 in methylase and at amino acid positions 155-159 in endonuclease. Lack of apparent similarity between methylase and endonuclease has also been observed for other systems, although the two enzymes recognize the

same target sequence. This is not surprising, for it has been demonstrated with EcoRI endonuclease that a small number of amino acid side chains in stereo-specific positions are involved in discrimination of the recognition sequence (16).

When the sequences of FokI enzymes were compared by computer with published sequences of other related enzymes, it was found that MFokI contained two copies of a segment of tetra-amino acids that has been identified as being characteristic of adenine methylases. The consensus sequence of the segment is Asp/Asn-Pro-Pro-Tyr/Phe (14,15), and it occurs once in E. coli dam methylase (78), EcoK hsdM gene product (15), T4 dam methylase (79), and adenine methylases of many other type II restriction-modification systems (EcoRI (58,59), EcoRV (61), PstI (62), DpnII (80), PaeR7I (64), CviBIII (81), TaqI (65)). This segment is not seen in any cytosine methylase (BspRI (82), BsuRI (63), DdeI (66), HhaI (83), EcoRII (84)). An interesting feature is that FokI methylase contained two copies of this segment, one at amino acid positions 218-221 and the other at amino acid positions 548-551. In contrast to many other adenine methylases that recognize symmetric sequences, the target sequences of FokI methylase are asymmetric. One interpretation is that each segment is responsible for methylation of different DNA strands. As reported in chapter VII, I found that this was indeed the case for FokI methylase. As indicated in Fig. 3, a low level of the modification phenotype was detected for pFokM1.3, in which about one-third of the C-terminal moiety of the MFokI gene has been deleted. This activity could be

due to the residual N-terminal domain.

SUMMARY

A DNA fragment that carried the genes coding for FokI endonuclease and methylase was cloned from the chromosomal DNA of F. okeanokoites, and the coding regions were assigned to the nucleotide sequence by deletion analysis. The methylase gene was 1,941 bp long, corresponding to a protein of 647 amino acid residues (Mr = 75,622), and the endonuclease gene was 1,749 bp long, corresponding to a protein of 583 amino acid residues (Mr = 66,216). The assignment of the methylase gene was further confirmed by analysis of the N-terminal amino acid sequence. The endonuclease gene was downstream from the methylase gene in the same orientation, separated by 69 bp. The promoter site, which could be recognized by E. coli RNA polymerase, was upstream from the methylase gene, and the sequences adhering to the ribosome-binding sequence were identified in front of the respective genes. Analysis of the gene products expressed in E. coli cells by gel filtration and SDS-PAGE indicated that the molecular weights of both enzymes coincided well with the values estimated from the nucleotide sequences, and that the monomeric forms were catalytically active. No significant similarity was found between the sequences of the two enzymes. Sequence comparison with other related enzymes indicated that FokI methylase contained two copies

of a segment of tetra-amino acids which is characteristic of adenine-specific methylase.

CHAPTER V

Overproduction and crystallization of FokI restriction endonuclease

RFokI, classified as a type IIS restriction endonuclease, recognizes an asymmetric pentanucleotide, and cleaves outside its recognition sequence as indicated by the arrows here: 5'-GGATG$(N)_9$↓-3' (68). Because of its unusual cleavage 3'-CCTAC$(N)_{13}$↑-5' specificity, some workers have suggested that it can be used differently from other type II enzymes. Unique 5'-protruding ends liberated by digestion with RFokI permit the efficient simultaneous ligation of several restriction fragments. Some kinds of vector were constructed for this purpose: gene synthesis and DNA segment replacement with the use of such vectors have been reported (85,86). RFokI cleaves at a set distance from the recognition sequence. With the use of RFokI and the adaptor-primer that contains the RFokI recognition sequence and the ssDNA moiety complementary to the target DNA, ssDNA and dsDNA can be cleaved at any desired position (6,87).

I have isolated the genes for the FokI restriction-modification system and determined the complete nucleotide sequence (chapter IV). The system forms an operon in which RFokI gene is downstream from the MFokI gene. The expression of the system is regulated by a promoter upstream of the MFokI gene. I

have constructed a plasmid, pFokMR5.2a, that carries the FokI restriction-modification gene inserted downstream of the lac promoter, and purified RFokI from E. coli cells carrying the plasmid. The active form of RFokI is monomeric and the molecular weight is 66,000. The amino acids of the N-terminal have not yet been sequenced.

Here, I constructed an overproducing strain and purified enough RFokI to crystallize it. The N-terminal amino acid sequence of RFokI was analysed.

MATERIALS AND METHODS

Plasmids and strains The E. coli host strains, JM109 (70) and UT481 [lon Δ(lac-pro) thyA met supD hsdR hsdM/F'traD36 proAB $lacI^q$ ZΔM15] (88), were from the collection of the laboratory of Takara Shuzo Co., Ltd. Plasmids were obtained as follows: pKK223-3 (carries a tac promoter) was from Pharmacia P-L Biochemicals, and pACYC184 (89,90) and pFokMR5.2a (carries genes for FokI restriction and modification (chapter IV)) were from the collection of the laboratory of Takara Shuzo Co., Ltd. Plasmids pFokMR4.3, pFokM2.9, pFokR2.35, and pFokR2.2 were constructed from pFokMR5.2a for this experiment. Their structures are given in Fig. 1.

Enzymes and chemicals Restriction endonucleases and T4 DNA ligase were products of Takara Shuzo Co., Ltd., and were used

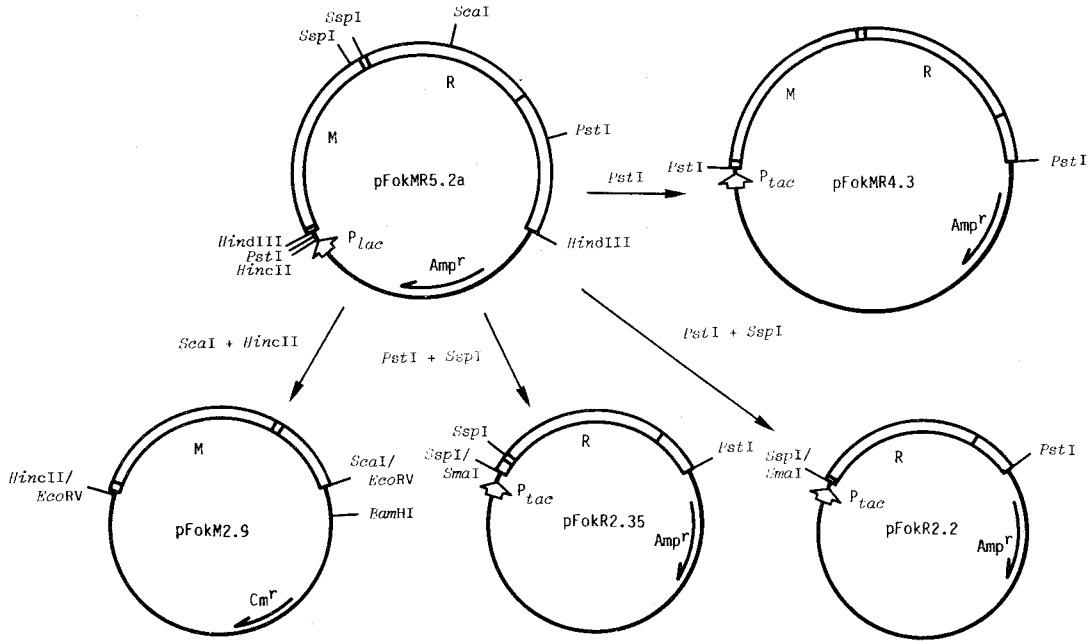

Figure 1. Structures of plasmids. pFokMR4.3 was constructed by insertion of the 4.3-kb PstI fragment that carried the genes for FokI methylase and endonuclease into the PstI site of pKK223-3. pFokM2.9 carried the FokI methylase gene inserted into the EcoRV site of pACYC184. pFokR2.35 and pFokR2.2 were constructed by replacement of a small SmaI-PstI fragment of pKK223-3 by the SspI-PstI fragments (2.35 kb and 2.2 kb, respectively) carrying the FokI endonuclease gene. The open box and the line show the insert and the vector moiety, respectively. Symbols: Amp^r, β-lactamase gene; Cm^r, chloramphenicol acetyltransferase gene; R, FokI endonuclease gene; M, FokI methylase gene; P_{lac}, lac promoter; P_{tac}, tac promoter.

according to the manufacturer's instructions. Protein standards for SDS-PAGE, the protein assays, and isoelectric focusing were from Bio-Rad and SDS-polyacrylamide gradient gels were from Daiichi Pure Chemicals.

Enzyme activity RFokI was diluted with 10 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerol, and its activity was assayed by incubation for 1 h at 37°C in 50 µl of a reaction mixture containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 60 mM NaCl, 0.01% BSA, and 1 µg of λ-DNA. One unit is defined as the amount of enzyme that cleaves 1 µg of λ-DNA in 1 h under these conditions.

Growth of cells For protein analysis of the E. coli cells carrying overproducing plasmids, cells were grown at 37°C in 50 ml of LB-medium (1% tryptone, 0.5% yeast extract, and 1% NaCl, pH 7.4) containing 100 µg/ml ampicillin, 30 µg/ml chloramphenicol, or both. When the cell density reached A_{600} = 0.6, IPTG was added to the concentration of 2 mM and incubation was continued at 37°C for 3 h.

For purification of RFokI, cells of E. coli UT481[pFokM2.9, pFokR2.2] were grown aerobically at 37°C in 40 l of LB-medium containing 100 µg/ml ampicillin and 30 µg/ml chloramphenicol. When the cell density reached A_{600} = 0.6, IPTG was added to 0.2 mM. After 3 h of induction at 37°C, cells (120 g) were harvested, washed with 0.15 M NaCl in 10 mM Tris-HCl, pH 7.5, and stored frozen.

Purification of RFokI From 120 g of cells, RFokI was purified as described in chapter IV. The cell-free extract was treated with polyethyleneimine and chromatographed on a phospho-cellulose and then a DEAE-cellulose column. The activity, recovered in the flow-through fraction, was adsorbed to a

hydroxylapatite column (Clarkson) and eluted with a linear gradient of potassium phosphate buffer, pH 7.5 (0.1 to 0.5 M). The active fraction obtained was further fractionated on a heparin-Sepharose (Pharmacia CL-6B) column and eluted with a linear gradient of KCl (0 to 1 M). The activity was eluted at 0.22-0.32 M KCl.

Analysis of N-terminal amino acid sequences RFokI from E. coli JM109[pFokMR5.2a] was blotted from SDS-polyacrylamide gel to Polyvinylidene difluoride membranes (Millipore) by the method of Matsudaira (75). RFokI from both F. okeanokoites and E. coli UT481[pFokM2.9, pFokR2.2] was desalted and analysed. The protein was degraded sequentially with a protein analyzer (Applied Biosystems model 470A) equipped with an on-line HPLC apparatus (model 120A).

Isoelectric focusing Isoelectric focusing was performed at 15°C on 1% agarose gels containing 4% ampholytes (Bio-Rad, Biolyte 8/10) in a Resolute HMP Chamber (FMC Corp.) by the procedure recommended by this manufacturer. The enzyme was desalted and samples of about 18 μg were put onto the gel with marker proteins. The samples were focused at a constant voltage of 500 V for 2 h. The gel containing the marker proteins and samples was fixed with trichloroacetic acid and stained with Coomassie brilliant blue. The remainder of the gel was cut into 30 segments. Slices of gels were placed in 2 ml of distilled water overnight, and the pH of the water was measured at 15°C.

Other methods Protein was assayed by the method of

Bradford (91) with use of a kit from Bio-Rad. SDS-PAGE was done by the method of Laemmli (92). DNA ligation, transformation, and plasmid isolation were done as described by Maniatis et al. (74).

RESULTS

Construction of plasmids overproducing RFokI The addition of an inducer only slightly increased the production of RFokI in cells of JM109[pFokMR5.2a], in which FokI methylase and endonuclease genes were under the control of the lac promoter (Fig. 2, lanes 1 and 2). To overproduce RFokI, the plasmid pFokMR4.3 was constructed by insertion of the genes for FokI methylase and endonuclease downstream from the DNA fragment containing the tac promoter. Production of methylase increased after induction, but the yield of endonuclease increased slightly (Fig. 2, lanes 3 and 4). Analysis of the nucleotide sequence of the FokI restriction-modification gene showed that there was a stem-loop structure followed by a T cluster between the end of the methylase gene and the beginning of the endonuclease gene (Fig. 3). I suspected that this putative translational termination structure interfered with the initiation of translation of the endonuclease, so a plasmid deficient in that structure was constructed.

I devised a two-plasmid system in E. coli with use of the compatible plasmids pKK223-3 and pACYC184. Plasmid pFokM2.9

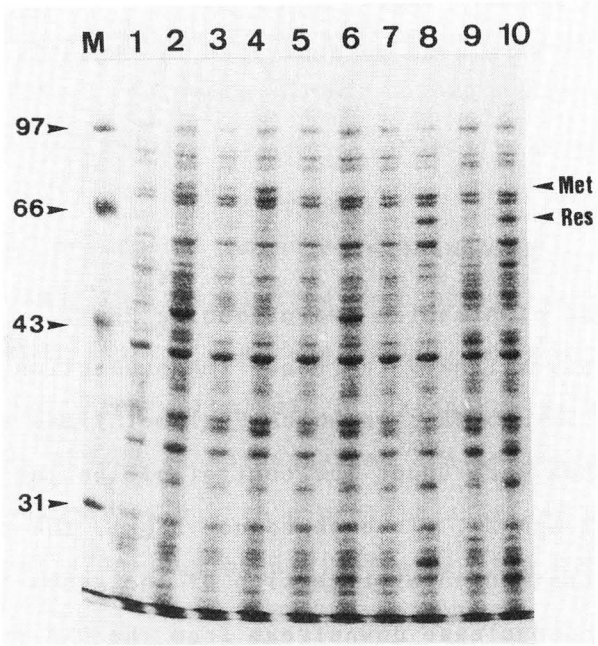

Figure 2. Overexpression of RFokI in E. coli cells. Cells carrying overproducing plasmids were grown as described in the text. Total proteins were extracted from cells harvested before induction (lanes 1, 3, 5, 7, 9) and after (lanes 2, 4, 6, 8, 10). Extract from an equivalent number of cells was put onto a lane, separated by electrophoresis on SDS-polyacrylamide gel (10%), and stained with Coomassie brilliant blue. By the side of the marker (M) lane, the molecular weights (x 10^{-3}) of the markers are given. The protein bands corresponding to the FokI methylase (Met) and the FokI endonuclease (Res) are indicated by arrows.

carries the FokI methylase gene inserted into the EcoRV site of plasmid pACYC184. The methylase was expressed constitutively under the control of its own promoter. pFokR2.2 carries the endonuclease gene without the stem-loop structure, and pFokR2.35 carries the endonuclease gene and the complete stem-loop structure under the control of the tac promoter. The plasmids constructed were introduced into E. coli cells, proteins were separated by

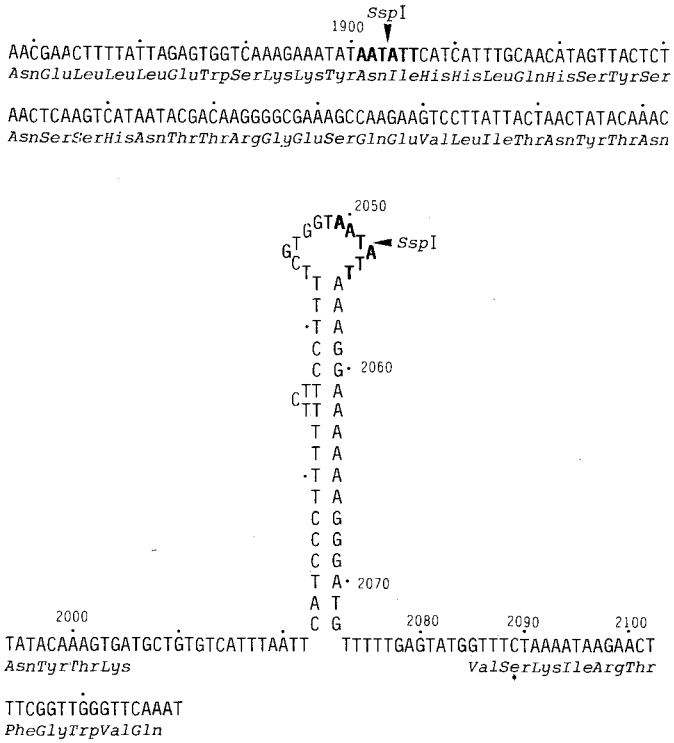

Figure 3. Nucleotide sequence and putative secondary structure upstream of the RFokI gene. Nucleotide number was the same as indicated in chapter IV. Two SspI sites, which was used for generation of the DNA fragment containing the RFokI gene, indicated as bold characters.

SDS-PAGE, and the RFokI activity was measured. JM109[pFokM2.9, pFokR2.35] production was on the same level as that of JM109[pFokMR5.2a] and JM109[pFokMR4.3]; the production of endonuclease increased greatly in cells carrying pFokM2.9 and pFokR2.2 (Fig. 2, lanes 5-10). RFokI activity was compare in E. coli strains carrying pFokMR5.2a, pFokMR4.3, [pFokM2.9, pFokR2.35], or [pFokM2.9, pFokR2.2]. The activity was stimulated 9-fold in JM109[pFokM2.9, pFokR2.2] and 30-fold in UT481

[pFokM2.9, pFokR2.2] cells.

Induction of RFokI activity The effect of the concentrations of IPTG on the induction of RFokI activity was examined. UT481[pFokM2.9, pFokR2.2] cells were grown to A_{600} = 0.6 and treated with 0.2 or 2 mM IPTG. After induction, some of the cells were harvested every hour and their endonuclease production was analysed by SDS-PAGE. With either concentration of IPTG, production was maximum after 3 h of induction, and there was little difference in the level of production with different concentrations. The growth of cells carrying both pFokM2.9 and pFokR2.2 was compared with the growth of those carrying no plasmid. Inhibition was slight in the recombinant cells. JM109 [pFokR2.2] cells without pFokM2.9 were not viable.

Purification and crystallization of RFokI Table I summarizes the purification scheme by which 300 mg of RFokI was purified from 120 g of UT481[pFokM2.9, pFokR2.2] cells. Figure 4 shows the results of SDS-PAGE. The purified sample gave a single protein band. This fraction was dialyzed against distilled water, and put on a thin-layer agarose gel; then isoelectric focusing was done. A single band was observed and the pI value of the RFokI was 8.9 (Fig. 5). The purified enzyme was dissolved in 10 mM Tris-HCl, pH 7.5, containing 100 mM KCl, 10 mM 2-mercaptoethanol, and 0.1 mM EDTA, and then ammonium sulfate powdered finely was gradually added to the enzyme solution (10.0 mg/ml) until a faint turbidity was obtained. The concentration of ammonium sulfate at the end was about 50% saturation. The pH of the solution was kept

Table I Purification of R*Fok*I

Steps	Total protein (mg)	Total activity (10^4U)	Specific activity (U/mg)	Yield (%)
Crude extract	12,320	120	97	100
Phosphocellulose	498	99	1,988	83
DEAE-cellulose	234	50	2,137	42
Hydroxylapatite	313	149	4,744	124
Heparin-Sepharose	299	160	5,619	140

Protein was assayed as described in the text with BSA as the standard.

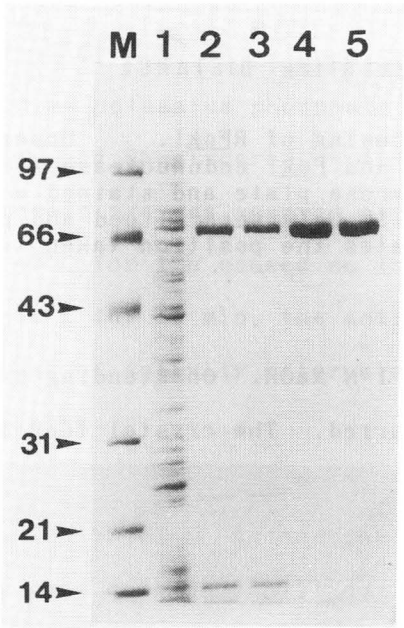

Figure 4. SDS-polyacrylamide gel electrophoresis of R*Fok*I. Proteins were electrophoresed on a SDS-polyacrylamide gradient gel(10% to 20%) and stained with Coomassie brilliant blue. Lanes 1 to 5 correspond to fractions of crude extract, phosphocellulose, DEAE-cellulose, hydroxylapatite, and heparin-Sepharose, respectively. By the side of the marker (M) lane, the molecular weight (x 10^{-3}) of markers are given.

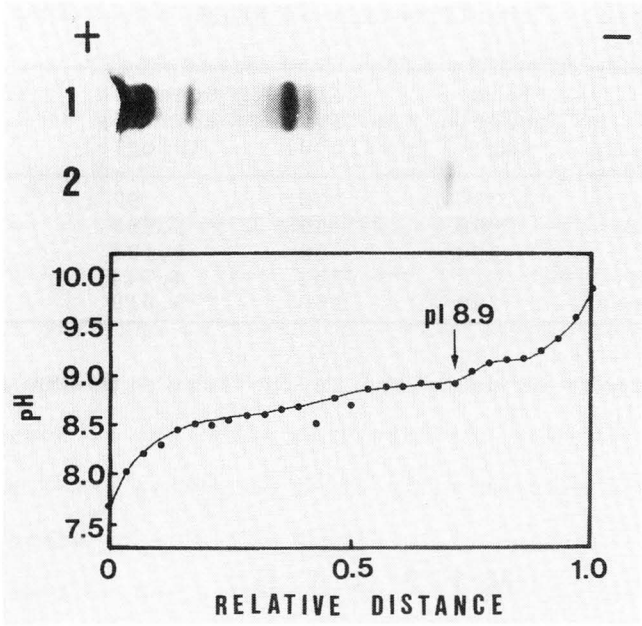

Figure 5. Isoelectric focusing of RFokI. Upper panel: Standard proteins (lane 1) and FokI endonuclease (lane 2) were focused on a thin-layer agarose plate and stained with Coomassie brilliant blue. Lower panel: Gels were sliced and pH values were measured. The arrow indicates the position taken to be the isoelectric pH.

constant at about 7.2 with 1 N NaOH. On standing overnight at 4°C, crystal formation occurred. The crystal took a form of thin plates (Fig. 6).

N-terminal amino acid sequencing The first eight amino acids of RFokI at the final purification step were (Met,Val)-(Ser,Phe)-(Leu,Lys)-(Ile,Ser)-(Arg,Met)-(Thr,Val)-(Ser,Phe)-(Lys,Gly). This corresponded to the amino acid sequence, Met-Phe-Leu-Ser-Met-Val-Ser-Lys, that started at the nucleotide position 2070, and to the amino acid sequence, Val-Ser-Lys-Ile-Arg-Thr-Phe-Gly, that started at 2082; the ratio of the production of the two

proteins was 1:2-3. The N-terminal amino acid sequence of RFokI purified from E. coli JM109[pFokMR5.2a] and from F. okeanokoites was also analysed. The first ten amino acids were X-X-X-Ile-Arg-Thr-Phe-Gly-X-Val (X not identified) and Val-Ser-Lys-Ile-Arg-Thr-Phe-Gly-Trp-Val, respectively. These corresponded to the amino acid sequence that started at nucleotide position 2082.

Properties of the enzyme Of the conditions tested, enzyme activity was maximum at 37°C in 10 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 60 mM NaCl, and 0.01% BSA. The addition of 7 mM 2-mercaptoethanol or 0.01% Triton X-100 did not increase the activity. Without BSA, activity decreased to 40%. Activity was greatest at 37-42°C. The pH optimum was 7.5-8.5. RFokI was incubated in 10 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerol at different temperatures, and the activity remainings was assayed at different times. Incubation at 42°C for 1 h caused no loss of activity, but after incubation at 50°C for 10 min, the activity decreased to 3%. There was no loss of activity after storage at 20°C for 1 week or at 4°C for 2 months.

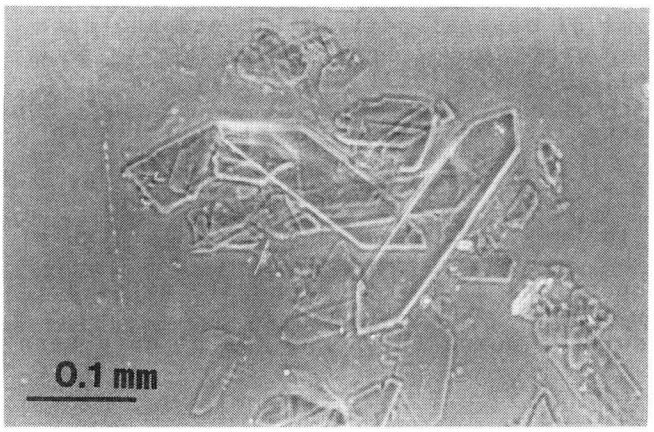

Figure 6. Crystals of RFokI.

DISCUSSION

I assumed that the weak translation of the endonuclease gene in the cloned FokI restriction-modification gene was caused by the stem-loop structure upstream of the endonuclease gene. Weak translation caused by mRNA folding has been reported for other restriction-modification systems. A stem-loop structure downstream of the endonuclease gene seems to interfere with EcoRV translation initiation (61,93). A 14 bp hairpin structure within the endonuclease coding sequence probably attenuates TaqI transcription or translation (94). In the TaqI system, overproduction was greatest when the hairpin structure was removed. Here, removal of the stem-loop structure increased the translation efficiency of the RFokI gene, and the RFokI gene was expressed efficiently under the control of the tac promoter.

Possible explanations for the high yield (140%) by my purification for RFokI include the use of the lon$^-$ E. coli strain as the host cells and the removal of substances that interfere with enzyme activity by hydroxylapatite chromatography. The pI of RFokI (8.9) was the highest of several restriction endonuclease (EcoRI, 6.4 (95), 6.3 (96); RsrI 7.0 (95); and AatII, 5.5 (97)). The high pI explains the recovery of RFokI activity in the pass-through fraction of DEAE-cellulose. An abundance of highly basic amino acids in RFokI might account for this high pI, so I investigated the proportions of basic and acidic amino acids. EcoRI, for which the amino acid composition and the pI have been

published, was studied in the same way. In RFokI, 16.5% of the amino acid residues were acidic and 12.4% were basic; in EcoRI, these figures were 14.8% and 12.6%, respectively. The difference was not significant.

The purified enzyme gave a single band on SDS-PAGE and isoelectric focusing gel electrophoresis, but analysis of the N-terminal amino acid sequence showed that the sample was a mixture of two proteins. One started at valine and the other contained an additional five amino acids at the N-terminal. The N-terminal amino acid of the endonucleases purified from F. okeanokoites and from E. coli JM109[pFokMR5.2a] was valine. The protein that has an additional five amino acids in the N-terminal and that was produced in the overproducing strain was probably an artifact of the gene manipulation. The following could explain the production of the artifact. In the FokI restriction-modification operon RFokI is translated from ATG at 2082 and is processed to become mature RFokI by removal of the methionine at the N-terminal. Enzymatic activity that removes the N-terminal residue of methionine when it precedes valine has been observed in both prokaryotes and eukaryotes (98). In the overproducing strain, a stable secondary structure around 2070 is destroyed and the ATG at 2070 can be used as an initiation site in addition to the ATG at 2082. Initiation codons at 2070 and 2082 are accompanied by putative ribosomal-binding sequences as follows: AAAGGAAAAAGGGATG, at 2070; and AAGGGATGTTTTTGAGTATG, at 2082 (the underlined nucleotides are complementary to 16s-rRNA). It is reasonable to

think that in pFokR2.2, translation starts from the initiation codon at either 2070 or 2082. Although the protein translated from ATG at 2082 is processed as described above, newly synthesized protein is not processed and thus accumulates. The results obtained here suggest that the insertion of a strong promoter just upstream of ATG at 2082 would bring about overproduction of the homogeneous RFokI protein that had valine at the N-terminal.

It was not known whether the protein that had an additional five amino acids at the N-terminal of RFokI had the same catalytic properties as RFokI or not. To check this, it would be necessary to separate the two proteins, but such separation would be difficult, because the differences in the molecular weights and isoelectric points are small.

Several restriction endonucleases were purified and crystallized from genetically engineered overproducing strains (17,61,94,99-103). Isolation of several hundred milligrams of RFokI will make possible its physical characterization by X-ray diffraction.

SUMMARY

To overproduce RFokI in an E. coli system, the coding region of RFokI predicted from the nucleotide sequence was generated from the FokI operon and joined to the tac promoter of an expression

vector, pKK223-3. By introduction of the plasmid into E. coli UT481 cells expressing the FokI methylase gene, the RFokI activity was overproduced about 30-fold, from which RFokI was purified in amounts sufficient for crystallization. The removal of a stem-loop structure immediately upstream of the RFokI coding region was essential for overproduction.

CHAPTER VI

Identification of nucleotides methylated by FokI methylase and characterization of the enzyme

The restriction-modification system found in F. okeanokoites recognizes an asymmetric DNA segment, 5'-GGATG-3' (68). MFokI
3'-CCTAC-5'
modifies DNA, and renders the DNA resistant to digestion by RFokI. Methylation of DNA by a site-specific methylase makes it possible to increase the cleavage specificity of restriction endonucleases (10). In the combination of MMspI and MFokI, RFokI specificity is increased from 5 (GGATG) to 7 (CCGGATG) bp (104). I have isolated the gene for the FokI restriction-modification system and sequenced its nucleotides (chapter IV). The amino acid sequence deduced from the nucleotide sequence shows that MFokI contains two domains characteristic of adenine methylase. This suggested that MFokI methylates adenine in the recognition sequence. Pósfai and Szybalski have shown that MFokIA methylates adenine, resulting in GGmATG (105). They suggest that another methylase is present in
CC TAC
vivo that methylates the bottom strand. Complete MFokI protein is responsible for the methylation of both strands in the target sequence (chapter VII). MFokI has been purified as a single protein from E. coli cells harboring a plasmid carrying the MFokI gene (chapter IV). The active form of MFokI is monomeric and its

molecular weight is 72,000. Here, I report on the reaction conditions and substrate specificity of MFokI and show evidence that MFokI methylates adenine in the recognition sequence.

MATERIALS AND METHODS

Plasmid and strain E. coli UT481 cells (88) were from the collection of the laboratory of Takara Shuzo Co., Ltd. The MFokI-overproducing plasmid pKS57, in which the MFokI gene is joined to the tac promoter, was described in chapter VII.

Enzymes and DNA Restriction endonucleases, T4 polynucleotide kinase, T4 DNA ligase, pUC18, and λ-DNA were products of Takara Shuzo Co., Ltd.

Other chemicals Protein standards for SDS-PAGE and the protein assays were from Bio-Rad, and SDS-polyacrylamide gradient gels were from Daiichi Pure Chemicals.

Measurement of MFokI activity Non-radioactive method MFokI activity was measured as described in chapter IV during the purification of the enzyme. To identify suitable reaction conditions, the method was slightly modified as follows. After incubation with methylase, the reaction was stopped by the addition of phenol. DNA was recovered by ethanol precipitation, suspended in 50 μl of 10 mM Tris-HCl, pH 7.5, 60 mM NaCl, 7mM $MgCl_2$, and 0.01% BSA, and incubated with 6 units of RFokI at 37°C for 1 h. One unit was defined as the amount of enzyme that

protected 1 μg of λ-DNA from cleavage by RFokI at 37°C for 1 h.

Radioactive method The enzyme was incubated in 20 μl of reaction mixture containing 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM 2-mercaptoethanol, 1 μCi of S-adenosyl-L-[methyl-^{3}H]methionine (Amersham, 85 μCi/mmol), and DNA, at 37°C. Then 15 μl of the mixture was spotted onto DEAE 81 paper (Whatman). The discs were rinsed in Na_2HPO_4 three times, washed with ethanol, and dried before ^{3}H was counted within a liquid scintillator.

Purification of enzyme E. coli UT481 cells carrying pKS57 were grown at 37°C in 20 l of LB-broth (1% tryptone, 0.5% yeast extract, and 1% NaCl, pH 7.4) containing 100 μg/ml ampicillin. When the cell density reached an A_{600} of 0.6, IPTG was added to the concentration of 0.2 mM. After 3 h of induction at 37°C, cells were harvested, washed with 0.15 M NaCl in 10 mM Tris-HCl, pH 7.5, and stored frozen. MFokI was purified as described in chapter IV. The cell-free extract was treated with polyethyleneimine and chromatographed on columns of phosphocellulose, DEAE-cellulose, and then hydroxylapatite. The active fractions obtained were further fractionated on a heparin-Sepharose (Pharmacia CL-6B) column and eluted with a linear gradient of KCl (0 to 1 M). The active fraction were concentrated and put on a Sephadex G-100 (Pharmacia) column. Active fractions eluted from Sephadex G-100 were concentrated and stored in 50% glycerol at -20°C.

Preparation of DNA The FokI linkers d(GGCCATCCGG) and d(CCGGATGGCC) were synthesized on DNA synthesizer (Applied

Biosystems 380A) and purified by reverse-phase HPLC. Phosphorylation and ligation of FokI linkers were done as described by Maniatis et al. (74).

Identification of the nucleotides methylated by MFokI Ligated FokI linker was labeled with ^{3}H as follows. First, 80 units of MFokI was added to a reaction mixture containing 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM 2-mercaptoethanol, 10 µCi ^{3}H-AdoMet, and ligated FokI linker with an OD of 0.01. Incubation was at 37°C for 30 min. After phenol extraction, 50 µg of sonicated calf thymus DNA (Sigma) was added as carrier and precipitated by ethanol. DNA was suspended in 100 µl of 100 mM sodium acetate, pH 5.3, and 0.5 mM $ZnCl_2$, and 20 µg of P1 nuclease (Yamasa) was added. After incubation at 37°C for 1 h, 40 µl of the reaction mixture was spotted onto a thin-layer plate (Avicel SF cellulose; Funakoshi). The plate was developed by the method of Nishimura (106). The positions of pdA, pdG, pdT, and pdC were identified by UV irradiation at 254 nm. The surface of the thin-layer plate was sprayed with EN^{3}HANCE (DuPont) and fluorography was done at -80°C overnight.

RESULTS AND DISCUSSION

Purification of MFokI Table I summarizes the method by which MFokI was purified from 77 g of UT481 cells carrying pKS57. Active fractions eluted from the heparin-Sepharose column at 0.10-

Table I Purification of MFokI

Steps	Total protein (mg)	Total activity (10^4U)	Specific activity (U/mg)	Yield (%)
Crude extract	7,050	2,430	3,500	100
Phosphocellulose	113	63.0	5,600	2.6
DEAE-cellulose	74.8	78.3	10,500	3.2
Hydroxylapatie	31.8	74.4	23,400	3.1
Heparin-Sepharose	0.66	16.0	242,400	0.7
Sephadex G-100	0.55	12.8	232,700	0.5

Protein was assayed as described in the text with BSA as the standard.

0.14 M KCl were analysed by SDS-PAGE. In the fractions that eluted at 0.10-0.11 M KCl, the 72-kDa protein was not observed but several proteins ranging from 40 to 43 kDa were. The fractions that eluted at 0.12-0.14 M KCl contained the 72-kDa protein as a major component. The activity level of MFokI that lacks its C-terminal one-third was as high as that of the 72-kDa MFokI protein, so the proteins corresponding to 40-43 kDa were probably products of degradation during the purification. I collected the fractions that eluted at 0.12-0.14 M KCl and further purified them on a Sephadex G-100 column. Figure 1 shows results of SDS-PAGE in the course of purification. The purified sample gave a single band at 72 kDa.

Properties of the enzyme Effects of pH, temperature, salts, and metals on the activity of MFokI were examined (Fig. 2). The addition of salt up to 50 mM had no effect on the activity,

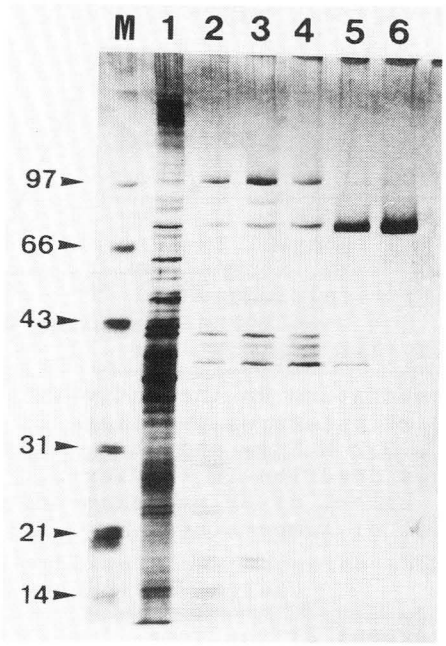

Figure 1. SDS-polyacrylamide gel electrophoresis of MFokI. Proteins were electrophoresed on a SDS-polyacrylamide gradient gel (10% to 20%) and silver-stained. Lanes 1 to 6 correspond to fractions of crude extract, phosphocellulose, DEAE-cellulose, hydroxylapatite, heparin-Sepharose, and Sephadex G-100, respectively. By the side of the marker (M) lane, the molecular weights ($\times 10^{-3}$) of the markers are given.

but higher concentrations were inhibitory (Fig. 2A). There was no difference between the action of NaCl and KCl. The optimum pH was 9.1, and the optimum temperature 37°C (Fig. 2B, C). Without EDTA, the activity decreased to 25%. The addition of $MgCl_2$ without EDTA further decreased the activity. The addition of 0.01% BSA or 0.01% Triton X-100 had no effect on the activity.

Substrate specificity The substrate specificity of MFokI is shown in Table II. DNA fragments cleaved with type II restriction endonucleases do not have the recognition site of the corresponding methylase, so the methyl group is not transferred to the restriction fragment. On the other hand, RFokI cleaves outside the recognition sequence, so the recognition site of MFokI is intact after cleavage with RFokI. The methyl group is probably

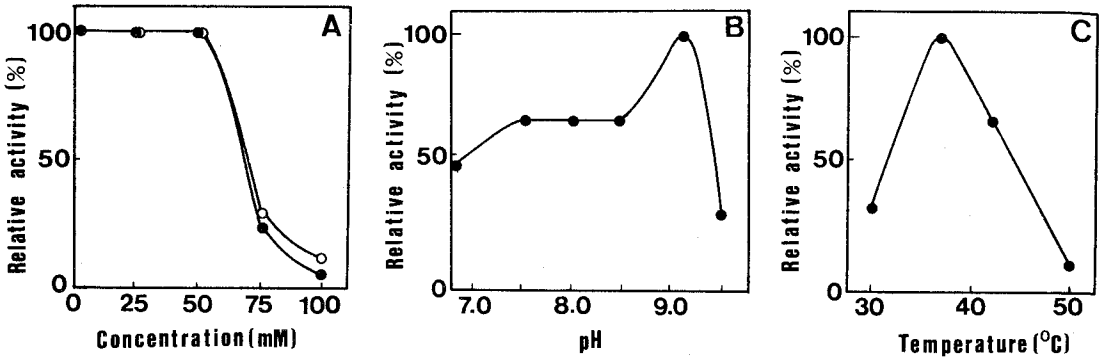

Figure 2. Effects of salt, pH, and temperature on the activity of MFokI. MFokI was diluted with 10 mM potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, 0.2 M KCl, and 5% glycerol, and its activity was assayed as described in the text. A, Effects of NaCl (●) and KCl (○). B, Effect of pH was examined with use of a Tris-HCl buffer. C, Effect of temperature.

transferred to the RFokI-cleaved DNA fragment at the same frequency as to the uncleaved DNA fragment. The amount of ^{3}H incorporated into the λ-FokI fragment DNA was comparable to that incorporated into the λ-DNA (Table II). With the 1117 bp pUC18-EcoRI,DraI fragment, in which there is no FokI recognition sequence, ^{3}H was not incorporated. These results suggest that MFokI recognized GGATG and methylated a nucleotide within the recognition sequence. To identify the nucleotide methylated by MFokI, I synthesized complementary oligonucleotides, d(GGCCATCCGG) and d(CCGGATGGCC), with no adenine except for the FokI recognition sequence. With ligated FokI linker, ^{3}H was incorporated at the same level as to λ-DNA. But with phosphorylated non-ligated FokI linker and non-phosphorylated linker, only 30% and 3% incorpotation was observed, respectively. These results indicate that ligated linker is as good a substrate as natural DNA.

Table II Substrate specificity of MFokI

Substrate	cpm	
	-MFokI	+MFokI
λ	225	192,711
λ-FokI fragment	184	149,207
pUC18-EcoRI,DraI fragment	304	4,509
Dephosphorylated FokI linker	208	370
Phosphorylated FokI linker	288	40,135
Ligated FokI linker	272	145,684

DNA was incubated with 8 units of MFokI for 30 min in the reaction mixture described in the text. Amounts of each substrate were as follows: λ, λ-FokI, and the pUC18-EcoRI,DraI fragment were each 1 μg, dephosphorylated, phosphorylated, and ligated FokI linker were all 0.002 OD. Except for pUC18-EcoRI,DraI, the amount of substrate was adjusted so as to contain equal number of FokI recognition sites.

Identification of nucleotide methylated by MFokI Ligated FokI linker was labeled with ^{3}H and then degraded to 5'-mononucleotides by P1 nuclease. The products were separated by two-dimensional thin-layer chromatography and the ^{3}H-labeled nucleotides were detected by fluorography (Fig. 3). A radioactive spot appeared at the position corresponding to N^{6}-methyladenosine 5'-monophosphate. No other labeled products were found with this substrate. Landry et al. purified MFokI from F. okeanokoites and found that it methylates the adenine residue in both strands of the recognition sequence (107). The results obtained by the use of MFokI purified from E. coli cells carrying the MFokI gene were the same as theirs. The results indicate that MFokI transferred a methyl group to the adenine in the recognition sequence. The

results shown in this chapter and chapter VII showed that M*Fok*I recognized GGATG and methylated adenine in both strands.
CCTAC

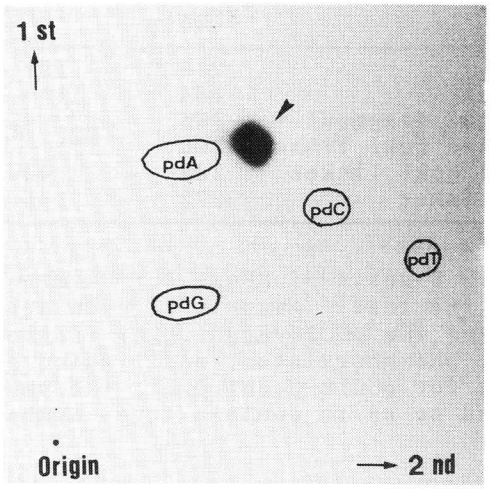

Figure 3. Fluorography of two-dimensional thin-layer chromatography. The plate was developed at room temperature as follows: first dimension, isobutyric acid : 0.5 M NH_4OH = 5 : 3 (v/v); second dimension, isopropanol : conc. HCl : H_2O = 70 : 15 : 15 (v/v/v). The positions of pdA, pdT, pdG, and pdC were detected by UV irradiation at 254 nm. The plate was sprayed with EN^{3}HANCE and fluorography was done at -80°C overnight. The arrow shows the position of N^6-methyladenosine 5'-monophosphate.

SUMMARY

From *E. coli* cells carrying an overproducing plasmid, M*Fok*I was purified to homogeneity. Unlike R*Fok*I, M*Fok*I requires EDTA and is inhibited by Mg^{2+}. With use of synthetic oligonucleotides, the sequence specificity was determined by analysis of a ^3H -

labeled nucleotide. M<u>Fok</u>I recognized 5'-GGATG-3' and methylated
3'-CCTAC-5'

adenine on both strands.

CHAPTER VII

Presence of two domains in FokI methylase for modification of different DNA strands

The FokI restriction-modification system found in F. okeanokoites recognizes a five bp segment of DNA, consisting of 5'-GGATG-3' in one strand and 3'-CCTAC-5' in the other (FokI recognition site), and the RFokI introduces staggered cleavages 9 and 13 nucleotides downstream from the recognition site (68). The genes coding for the FokI system have been cloned, and their entire nucleotide sequences have been given in chapter IV. The assigned structural gene for the MFokI was 1,941 bp long, encoding a protein of 647 amino acid residues. Comparison of the amino acid sequence with the sequences of other methylases showed that MFokI contains two copies of a segment of tetra-amino acids, which is characteristic of adenine-specific methylases (14,15). The consensus sequence for this segment is Asp/Asn-Pro-Pro-Tyr/Phe, and MFokI has Asp-Pro-Pro-Tyr at amino acid positions 218 to 221 (segment 218-221) and 548 to 551 (segment 548-551). Therefore, MFokI is most likely to be an adenine-specific methylase. However, this adenine methylase-specific segment occurs only once in other methylases that belong to the type II restriction-modification systems (MEcoRI (58,59), MEcoRV (61), MPstI (62), MDpnII (80), MPaeR7I (64), MCviIII (81), MTaqI (65), and in the dam

methylases of E. coli (78) and T4 phage (79)). These methylases differ from MFokI in that their target sequences are all symmetrical.

In the deletion analysis described in chapter IV, some MFokI activity was conferred on E. coli cells by plasmid pFokM1.3 carrying a truncated MFokI gene, although the activity level was not high as that conferred by the wild-type gene. The gene in pFokM1.3 lacks in the C-terminal one-third, which includes the segment 548-551. A simple interpretation of the result is that MFokI has two functionally identical domains for methylation, and that the full activity is retained by their both acting. Another possibility is that because the MFokI modifies the asymmetric sequences, the two domains carrying each segment may participate in methylation of adenine residues in different strands of the target DNA. To explore these possibilities, I introduced mutations into each of the two segments by oligonuclotide-directed mutagenesis, and analysed the activity levels and the methylated strands. Different DNA strands were modified asymmetrically by introduction of mutations.

MATERIALS AND METHODS

Strains E. coli CJ236 [dut1 ung1 thi1 relA1/pCJ105(Cm^r)] was used for preparation of M13 viral DNA containing uracil (108). E. coli JM109 was used as a host for most of the plasmid

constructs derived from pUC19, M13mp18, and M13mp19 (70). JM109 was also used as a *lacI*^{q} host for the expression vector pKK223-3 (Pharmacia P-L Biochemicals), which uses the *tac* promoter for expression of cloned genes (Fig. 1).

Enzymes and biochemicals The restriction enzymes, bacterial alkaline phosphatase, T4 DNA ligase, T4 DNA polymerase, T4 polynucleotide kinase and M13 sequencing kit were donated by Takara Shuzo Co., Ltd., and used according to the manufacturer's instructions. AdoMet was obtained from Sigma Chemical Co. [α-^{32}P]dCTP for DNA sequencing and [γ-^{32}P]ATP for end-labeling were from Amersham and New England Nuclear Corp., respectively. The calibration kit for molecular weights of protein and SDS-polyacrylamide gradient gels were products of Pharmacia P-L Biochemicals and Daiichi Pure Chemicals, respectively.

Oligonucleotide-directed mutagenesis The 36-mer, d(CCCATTATAAGGGGGACCTATGTATAAAATATCCC), and the 35-mer, d(TAATTAAATAGGGTGGTGCACAGTATACTAAATC), were synthesized by a Beckman System-1 Plus DNA synthesizer. Each of these oligonucleotides carries two mismatched nucleotides, and creates both an amino acid substitution and a new restriction site at the target region (Fig. 2A). After purification by polyacrylamide gel electrophoresis, their 5'-termini were phosphorylated in polynucleotide kinase reaction. M13 viral DNA that carried the entire M*Fok*I gene (Fig. 1) was propagated in *E. coli* CJ236. To this viral DNA containing uracil, synthesized oligonucleotides were hybridized and extended with the use of T4 DNA polymerase by

the protocol of Kunkel et al. (108). T4 DNA ligase was used to close the strands, and the products were introduced into E. coli JM109 cells. The resulting plaques were isolated, and mutants were screened by analysis of newly generated restriction sites. The mutation sites were confirmed by analysis of DNA sequences.

Induction of MFokI genes inserted in the expression vector Wild-type and mutant MFokI genes were inserted into the expression vector pKK223-3, and the constructs were introduced into E. coli JM109. The cells were grown at 37°C in M9 medium (0.6% Na_2HPO_4, 0.3% KH_2PO_4, 0.1% NH_4Cl, 0.05% NaCl, 1mM $MgSO_4$, 0.1 mM $CaCl_2$) supplemented with 0.5% glycerol, 0.001% thiamine, 0.4% casamino acids, 50 µg/ml ampicillin. When the cell density reached A_{600} of 0.25, IPTG was added to 2.5 mM. After 3 h of induction at 37°C, cells were harvested, washed with 0.14 M NaCl, 20 mM Tris-HCl, pH 7.5, and stored frozen.

Protein analysis The JM109 cells induced by IPTG were suspended in 0.0625 M Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, and 0.01% bromophenol blue, and after boiling for 5 min, solubilized proteins were electrophoresed on SDS-polyacrylamide gradient gels (10% to 20%) as described by Laemmli (92). Protein bands were made visible by staining with Coomassie brilliant blue R-250.

Purification of MFokI gene products About 5 g of JM109 cells induced by IPTG and carrying MFokI gene:pKK223-3 recombinants were thawed and suspended in 25 ml of 50 mM Tris-HCl, pH 7.5, 2 mM EDTA, 1 mM dithiothreitol, 0.1 mM

phenylmethylsulfonylfluoride, and 0.5 mg/ml lysozyme. After holding for 30 min on ice, the mixture was briefly sonicated to complete lysis, and centrifuged (10^5 x g for 1 h) to obtain the crude extract. The following steps of purification were identical to those described in chapter IV. The extract was treated with polyethyleneimine and chromatographed on phosphocellulose and DEAE-cellulose columns. The final products were concentrated by dialysis against 50 mM Tris-HCl, pH 7.5, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, and 50 % glycerol, and stored at -20°C. Under these conditions, activity was stable for at least 3 months.

Assay of methylase activity in vivo Plasmids were purified by CsCl-ethidium bromide centrifugation from JM109 cells induced by IPTG and carrying MFokI gene:pKK223-3 recombinants. One microgram of the plasmids was treated for 2 h at 37°C with 1 or 3 units of RFokI in 20 µl of reaction mixture containing 10 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 60 mM NaCl, and 7 mM 2-mercaptoethanol. The digests were electrophoresed on 1% agarose gel, and made visible by staining with ethidium bromide.

Identification of methylated strands Two picomoles of the 289 bp fragment carrying a single FokI recognition site (Fig. 5) were dephosphorylated by bacterial alkaline phosphatase, and rephosphorylated by the use of [γ-^{32}P]ATP (6000 Ci/mmol) and T4 polynucleotide kinase. The specific activity of the products was about 1.5 x 10^4 cpm/femtomole. One hundred femtomoles of the 5'-^{32}P-labeled DNA fragment were incubated with purified MFokI gene-products in 50 µl of reaction mixture containing 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM 2-mercaptoethanol, 0.1 mM AdoMet,

under the conditions that caused more than 99% of the dsDNA fragment to be rendered resistant to RFokI cleavage. Four femtomoles of the labeled DNA, treated or not-treated, were denatured by holding for 5 min at 100°C in 16 µl of reaction mixture containing 12.5 mM Tris-HCl, pH 7.5, 8.8 mM $MgCl_2$, and 75 mM NaCl. The reaction mixture was added 1 µl (400 fmol) of M13 viral DNA (M13KS24 ssDNA and M13KS26 ssDNA in Fig. 5). The mixtures were kept for 60 min at 65°C, and after annealing, 1 µl of 140 mM 2-mercaptoethanol, 1 µl (0.1 µg) of DNA and 2 µl (2 units) of RFokI were added. The mixture was incubated for 60 min at 37°C, and the product size was analysed by electrophoresis on 6% polyacrylamide gel under the denaturing conditions, followed by autoradiography.

Other procedures DNA ligation, transformation and plasmid isolation were done as described by Maniatis et al. (74). DNA was sequenced by the dideoxy-chain termination methods (76,77).

RESULTS

Construction of amino acid-substitution mutants Starting from pFokMR5.2a, which carried the complete MFokI and RFokI genes, the left 2.4 kb region (the BamHI site at the leftmost end to the ClaI site in Fig. 1), which encompasses the entire MFokI gene, was inserted into the BamHI-AccI region in the multicloning sites of

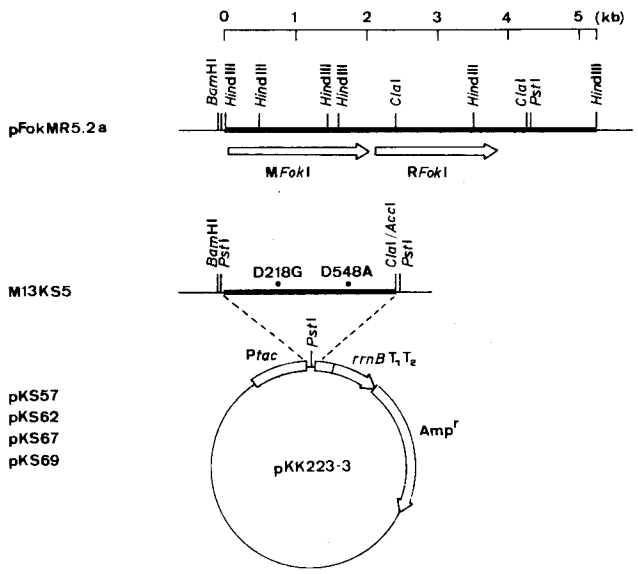

Figure 1. Construction of plasmids overproducing MFokI. The physical map of the 5.2 kb insert of pFokMR5.2a, spanning the MFokI and RFokI genes (open arrows), is shown at the top, on which the major restriction sites are indicated. The 2.4 kb region from the left end-BamHI site to the first ClaI site, which covered the MFokI gene, was first cloned in the BamHI and AccI sites of M13mp18. The 2.4 kb insert in the construct, M13KS5, was then generated by PstI, and cloned in the PstI site of the expression vector pKK223-3, so that the MFokI gene was placed under the control of the inducible tac promoter (Ptac) and rrnB terminator ($rrnBT_1T_2$). pKS57, pKS62, pKS67 and pKS69 are constructs carrying the genes for MFokIwt, MFokID218G, MFokID548A, and MFokID218G:D548A. Asterisks on the map of M13KS5 indicate the positions of D218G and D548A mutations.

M13mp18. To the resulting construct, named M13KS5, the synthesized 35-mer or 36-mer was hybridized, and oligonucleotide-directed mutagenesis was carried out. The sequences of these oligonucleotides were designed not only to replace a single amino acid but also to create a new restriction site (Eco0109I or ApaLI) in each of the Asp-Pro-Pro-Tyr segments (Fig. 2A). This made it easy to screen for the amino acid-substitution mutants, and two

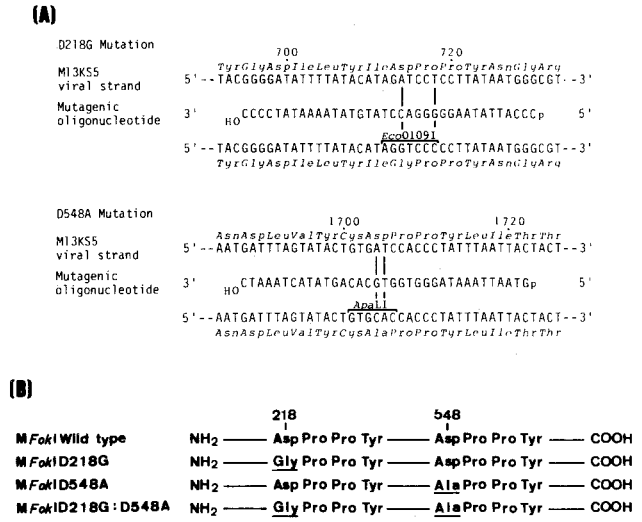

Figure 2. Construction of D218G and D548A mutations (A), and the structures of MFokI gene products (B). In (A), the nucleotide sequences of the viral strand of M13KS5 used for construction of D218G and D548A mutations, oligonucleotides used for mutagenesis, and nucleotide sequences obtained after mutagenesis are indicated, aligned in the same coordinates. The amino acid-sequences encoded thereby are given above and below the nucleotide sequences. Numbers above the nucleotide sequences are the positions from the leftmost end of the 5.2 kb insert in pFokMR5.2a. The mismatched nucleotides in the oligonucleotides were indicated by vertical lines, and newly created restriction sites by brackets. In (B), structures of the wild type and three mutants of MFokI, which were confirmed by nucleotide sequence analysis. The regions with identical sequences are represented by bars, except for the two adenine methylase-specific segments. Numbers are the amino acid positions from the N-termini, and substituted amino acids are underlined.

mutants were isolated, one with Asp replaced by Gly in segment. 218-221 and the other with Asp replaced by Ala in segment 548-551 These mutations were named D218G and D548A, respectively (see Fig. 1 and 2A). Then I constructed a double mutant, named D218G:D548A, which carried amino acid substitutions in both segments, using the two oligonucleotides for primer extension.

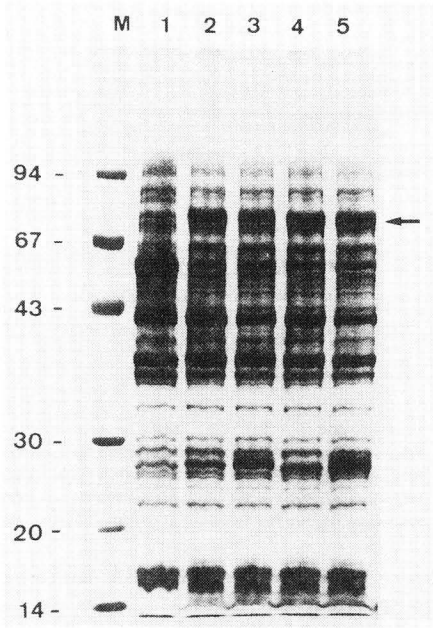

Figure 3. Overexpression of the wild-type and mutant MFokI genes in E. coli cells. E. coli JM109 cells carrying plasmids pKK223-3 as control (lane 1), pKS57 (lane 2), pKS62 (lane 3), pKS67 (lane 4), and pKS69 (lane 5) were grown in M9 media and induced by IPTG. The total protein was extracted from cells harvested at equivalent cell-densities, resolved by electrophoresis on SDS-polyacrylamide gradient gels (10% to 20%), and stained with Coomassie brilliant blue. By the side of the marker lane (lane M), molecular weights (x 10^{-3}) of markers are given.

Hereafter, MFokI of the wild type and its derivatives carrying mutations described above are abbreviated MFokIwt, MFokID218G, MFokID548A, and MFokID218G:D548A, respectively.

Overexpression of gene products First, the 2.4 kb fragments were generated from the M13 clones by PstI digestion, and were inserted into the PstI site of the expression vector, pKK223-3, in the proper orientation (Fig. 1). The resulting constructs, carrying MFokIwt, MFokID218G, MFokID548A, and

MFokID218G:D548A, were named pKS57, pKS62, pKS67, and pKS69, respectively. In each case, the authenticity of the alleles was confirmed by complete nucleotide sequencing. Thus, the MFokI alleles constructed should have the amino acid sequences shown in Fig. 2B. The plasmid constructs were introduced into E. coli cells, and after induction by IPTG, proteins expressed from the respective alleles were identified through SDS-PAGE. Compared to the cells harboring the vector alone (Fig. 3, lane 1), an intense band that migrated at the position of 75 kDa was observed for cells harboring pKS57, pKS62, pKS67 and pKS69 (the band with an arrow in Fig. 3, lanes 2 to 5). These proteins were judged to be MFokI gene products on the basis of the identity of their molecular weights with those of the purified methylase. Analysis also indicated that the products of mutant genes were as stable as MFokIwt regardless of these amino acid substitutions.

Methylase activities expressed in vivo of wild type and mutant genes The MFokI activities in E. coli cells, each of which carried one of plasmids pKS57, pKS62, pKS67 and pKS69, were compared in their susceptibility to RFokI of plasmids extracted from E. coli cells induced with IPTG (Fig. 4). pKS69 coding for MFokID218G:D548A was susceptible to RFokI to about the same extent as the control plasmid without methylation (Fig. 4, lanes 2 and 3), and was digested completely by incubation for 2 h with 3 units of RFokI per 1 μg of DNA (lanes 14 and 15). No digestion of pKS57 coding for MFokIwt occurred (lanes 5 and 6), and pKS62 and pKS67, which respectively coded for MFokID218G and MFokID548A, were

digested slightly only at increased RFokI concentrations (lanes 8 and 9 for pKS62; lanes 11 and 12 for pKS67). It is therefore obvious that MFokID218G and MFokID548A both retained significant levels of modification activity.

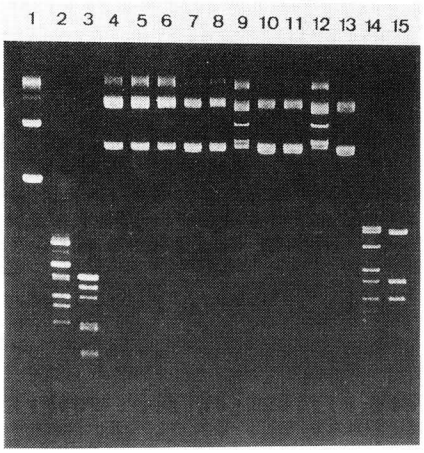

Figure 4. Assay of methylase activities expressed *in vivo* by the wild-type and mutant MFokI genes. Plasmids pKK223-3 as control (lanes 1-3), pKS57 (lanes 4-6), pKS62 (lanes 7-9), pKS67 (lanes 10-12), and pKS69 (lanes 13-15) were purified from JM109 cells treated with IPTG and that harbored each of these plasmids. One μg of plasmid was treated for 2 h with 1 unit (lanes 2, 5, 8, 11, 14) or 3 units (lanes 3, 6, 9, 12, 15) of RFokI, and analysed by electrophoresis on 1% agarose gels. Lanes 1, 4, 7, 10 and 13, are untreated controls.

Analysis of methylated strands by products of wild-type and mutant genes To gain insight into the mode of action of MFokI, the DNA strands methylated by MFokIwt, MFokID218G, or MFokID548A were analysed by the strategy shown in Fig. 5. In principle, both strands of a short dsDNA fragment carrying a single FokI recognition site near the middle were terminally labeled, and

separated by duplex formation with M13 viral DNAs in which the respective strands had been inserted. After digestion with RFokI, the sizes of terminally labeled fragments were measured. Modification in the hybridized strand can be detected from its susceptibility to RFokI.

The short dsDNA fragment used here was one generated from Col E1 by ClaI and StuI (nucleotide positions 2311 to 2566 in the ColE1 map (109)). This fragment was cloned into the multicloning sites of pUC19, and regenerated by EcoRI plus HindIII as a fragment of 289 bp. This made it easy to prepare M13 clones carrying the 289 bp fragment in either of two orientations by the use of M13mp18 and mp19. Viral DNAs prepared from these M13 clones were named M13KS24 ssDNA and M13KS26 ssDNA. Both strands of the 289 bp fragment were highly labeled with ^{32}P at the 5'-termini, and the products were reacted with purified MFokIwt, MFokID218G, or MFokID548A in the presence of AdoMet. After denaturation, they were hybridized with M13KS24 ssDNA and M13KS26 ssDNA. The duplexes were intensively digested with RFokI, and analysed by sequence gels under the denaturing conditions. The MFokI recognition site in the single stranded form is not susceptible to RFokI (110).

The results of analysis are shown in Fig. 6, in which lanes 1 to 4 are for untreated DNA fragments and lanes 5 to 8 are for those treated with MFokIwt, lanes 9 to 12 are for those treated with MFokID218G, and lanes 13 to 16 are for those treated with

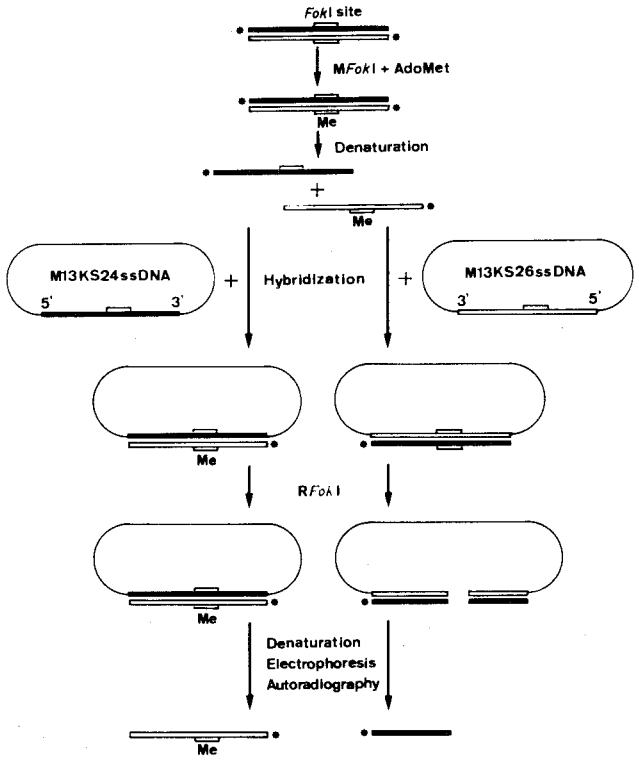

Figure 5. Strategy used for determination of methylated strands. The two strands of the 289 bp fragment, carrying the FokI recognition sequence (brackets), are terminally labeled (asterisks) and treated with MFokI gene products in the presence of AdoMet. The two strands are dissociated by heat-denaturation and hybridized with M13 viral DNA, carrying the respective strands of the 289 bp fragment (M13KS24 ssDNA and M13KS26 ssDNA). The duplexes are digested with RFokI, and after denaturation, the sizes of the labeled strands are measured. Only the non-methylated strand is shortened; the size of the methylated strand (indicated by "Me") is unchanged.

MFokID548A. With the untreated DNA fragments, the labeled fragment generated from the strand carrying 5'-GGATG-3' was 142 nucleotides long, and that generated from the other strand

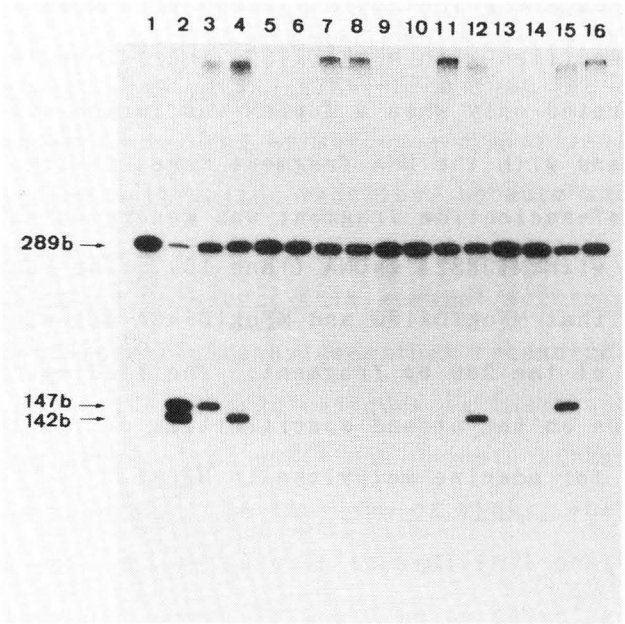

Figure 6. Analysis of methylated strands by the wild type and mutant MFokI gene products. Two femtomoles each of the 5'-^{32}P-labeled fragment (about 3 x 10^4 cpm) were treated with MFokIwt (lanes 5-8), MFokID218G (lanes 9-12) and MFokID548A (lanes 13-16). Lanes 1-4 are of the untreated control. These fragments were heat-denatured, and hybridized with M13KS24 ssDNA or M13KS26 ssDNA. The resulting duplexes were digested with RFokI, and the sizes of labeled strands were analysed by gel electrophoresis under denaturing conditions. In each set of four lanes, the first lane is the fragment alone (lanes 1, 5, 9, 13), the second lane is the fragment digested with RFokI without denaturation (lanes 2, 6, 10, 14), the third lane is the one, digested with RFokI after hybridization with M13KS24 ssDNA (lanes 3, 7, 11, 15), and the fourth lane is the one, digested with RFokI after hybridization with M13KS26 ssDNA (lanes 4, 8, 12, 16).

carrying 3'-CCTAC-5' was 147 nucleotides long (lanes 3 and 4). The generated fragments were the same sizes as those yielded from the original dsDNA fragment by RFokI digestion (lane 2). No cleavage of the 289 bp fragment that had been treated with MFokIwt was observed, regardless of duplex formation with M13 viral DNA

(lanes 6 to 8), indicating that both strands were modified. With the DNA fragment modified with MFokID218G, the 142-nucleotide fragment was generated only when a duplex was formed with M13KS26 ssDNA (lane 12), and with the DNA fragment treated with MFokID548A, the 147-nucleotide fragment was generated only when a duplex was formed with M13KS24 ssDNA (lane 15). The results clearly indicated that MFokID218G and MFokID548A methylate different strands of the 289 bp fragment. The findings also provide information on the strand specificities of the two segments specific for adenine methylase in MFokI.

DISCUSSION

Analysis showed that different DNA strands were modified asymmetrically by introduction of mutations into one of the two adenine methylase-specific segments. The wild-type enzyme modified both strands. The species and exact position of bases modified by MFokI were not identified, but it is known that methylases of the type II restriction-modification systems modify internal adenine or cytosine residues within the recognition sequences (111). It is therefore most likely that the bases attacked by MFokI are adenine residues within the FokI recognition site. The FokI enzymes recognize 5'-GGATG-3' in one strand and 3'-CCTAC-5' in the other, and the results obtained here indicated that the two domains, containing the segment 218-221 or 548-551,

participated in modification of the former and latter strands, respectively.

The result also provided evidence that the adenine methylase-specific segment, deduced by sequence comparison, was directly involved in the methylation reaction, because substitution of a single amino acid within the segment caused loss of activity. In contrast to M.FokI, however, this segment occurs only once in many other adenine-specific methylases that recognize symmetrical sequences. The two adenine residues in these sequences are in rotationally symmetrical positions on both strands of the recognition sequence. In the case of M.FokI, the molecule must attack two adenine residues in asymmetrical positions on dsDNA. For this type of enzyme, it would be an advantage for two different domains to be within the same molecule, each of which has a specific structure responsible for recognition of the strand-specific sequence and the common structure for methylation reaction.

Simultaneous modification of both strands would be a prerequisite for the chromosomal DNA of organisms, because nascent strands generated during DNA replication must be protected from the attack of site-specific endonucleases. It has been reported that the HsdM gene product in the type I restriction-modification systems contains a single copy of the adenine methylase-specific segment (15), whereas its target sequences are asymmetric. In this case, both strands at the target site are probably modified by some mechanism different from that operating in M.FokI. In the

paper by Pósfai and Szybalski (111), it was reported that with the MFokI preparation from New England Biolabs, only a unique adenine residue located on one of the strands was modified by their enzyme. Cloning of the MFokI gene has previously been reported from New England Biolabs (71). When the restriction map of their clone pDN204 was compared with that reported here, their gene was seen to be truncated at the HindIII site at nucleotide position 1434 in the map shown in Fig. 1. If the FokI methylase of New England Biolabs has been prepared from pDN204 or its derivatives, their enzyme lacks the segment 548-551, which may cause asymmetric modification.

SUMMARY

Based on the previous findings that the MFokI consists of 647 amino acid residues and contains two copies of the segment specific for adenine methylase, Asp-Pro-Pro-Tyr, at amino acid positions 218 to 221 and 548 to 551 (chapter IV), the role of these copies in the methylation reaction was investigated by introduction of a mutation into each segment. The MFokI gene was inserted into M13 vectors, and the Asp residues in the two segments were converted to Gly and Ala by oligonucleotide-directed mutagenesis. The wild-type and mutant genes were re-cloned into an expression vector, from which gene products were purified. A short DNA fragment carrying the FokI recognition site was treated

with each of these enzymes, and after separation of the two strands by duplex formation with M13 viral DNAs carrying the respective strands, the presence or absence of modification was judged from susceptibility to RFokI. The results of analysis showed that different strands were modified in an asymmetric way by the introduction of mutations into one of the two segments, and that the segments at the N-terminal and C-terminal moieties participated in modification of the strands carrying 5'-GGATG-3' and 3'-CCTAC-5', respectively. I concluded that MFokI contained two functional domains each of which was responsible for modification of different strands in the target DNA.

CONCLUSIONS

In this study, I dealt with restriction and modification enzymes produced in bacteria.

In chapter I, a method for the identification of the cleavage site of restriction enzymes was established. A synthetic oligonucleotide that contained the recognition sequence of the enzyme to be tested was used as the substrate. The cleavage site was deduced from the chain length of the products. Using this method, I determined the cleavage site of AccII.

During the identification of the cleavage site of several restriction enzymes by the method described in chapter I, it was appeared that the reactivity of the enzymes was dependent on the chain length of the substrate. In chapter II, the substrate specificity of the restriction enzyme ScaI was examined by the measurement of the kinetic constants for substrates of different chain lengths. Vmax values were similar but Km values were different. These findings gave useful information about the preparation of DNA for the identification of cleavage sites.

In chapter III, the purification and characterization of a new site-specific endonuclease, AccIII, were described. The enzyme was purified so as not to contain other nuclease or phosphatase activity. AccIII recognized 5'-TCCGGA-3' and cleaved between T and C. AccIII activity was inhibited by adenine methylation at the overlapping dam methylase recognition sequence. The optimum reaction conditions for AccIII were different from

those for other two restriction enzymes produced by the same organism. AccIII made it possible to cleave DNA at a unique site and is of potential usefulness for recombinant DNA experiments.

In chapters IV to VII, I described the restriction-modification system of Flavobacterium okeanokoites. In chapter IV, cloning and the nucleotide sequence of the gene were described. The methylase gene was 1,941 bp long, corresponding to a protein of 647 amino acid residues (Mr = 75,622), and the endonuclease gene was 1,737 bp long, corresponding to a protein of 579 amino acid residues (Mr = 65,737). The endonuclease gene was downstream from the methylase gene and in the same orientation, and the expression of both genes was regulated by a promoter in front of the methylase gene. Analysis of the gene products by gel filtration and SDS-PAGE showed that the molecular weight of both enzymes coincided well with the values estimated from the nucleotide sequences, and that the monomeric forms were catalytically active. Little similarity was found between the sequences of the two enzymes. Sequence comparison with other related enzymes indicated that FokI methylase contained two copies of a segment of tetra-amino acids characteristic of adenine methylase.

In chapter V, I reported on overproduction and characterization of FokI restriction enzyme. The plasmid in which the coding region of FokI restriction enzyme was joined to the tac promoter was introduced into E. coli cells expressing the FokI methylase gene, and the FokI restriction enzyme was overproduced.

FokI restriction enzyme was purified in amounts sufficient for crystallization. Removal of a stem-loop structure immediately upstream from the gene coding for FokI restriction enzyme was essential for overproduction.

Chapter IV described the purification and characterization of FokI methylase. From E. coli cells carrying an overproducing plasmid, FokI methylase was purified to homogeniety. EDTA was essential for the activity. FokI methylase recognized 5'-GGATG-3'
3'-CCTAC-5'
and methylated adenine within the recognition sequence.

Finally, I explained the role of amino acid domains of FokI methylase in the methylation reaction. FokI methylase contains two copies of the segment specific for adenine methylase, Asp-Pro-Pro-Tyr. The segments at the N-terminal and C-terminal moieties participated in modification of the strands carrying 5'-GGATG-3' and 3'-CCTAC-5', respectively.

ACKNOWLEDGMENTS

I express my deepest gratitude to Professor Hideaki Yamada, Kyoto University, for his kind guidance and encouragement throughout the course of this study.

I am also grateful to Professor Mituru Takanami, Kyoto University, for his valuable advice and warm encouragement in carrying out this work.

My sincere thanks are due to Dr. Hiroyuki Sugisaki, Research Associate of Kyoto University, without whose thoughtful advice and instructions this work would not have been possible.

It is a great pleasure to acknowledge Dr. Sakayu Shimizu, Research Associate of Kyoto University, for his encouragement and critical reading of the manuscript.

I am very much indebted to Mr. Hisashi Ohmiya, Senior managing director, Dr. Akira Obayashi, Director of Central Research Laboratories, and Dr. Teruya Nakamura, Director of the Bioproducts Development Center, Takara Shuzo Co., Ltd, for providing the chance to carry out these studies. I greatly appreciate Dr. Nobutsugu Hiraoka, Mr. Hirokazu Kotani, and other researchers at the Central Research Laboratories, Takara Shuzo Co., Ltd., for their support and many lively discussions.

I thank Professor Kazuo Yonaha, University of the Ryukyus, for his crystallization of the FokI restriction enzyme.

Last, but not least, I express my gratitude to Mr. Tsutomu Taniguchi, Takara Shuzo Co., Ltd., for his continuous

encouragement while I was preparing this thesis.

REFERENCES

1. Meselson, M., and Yuan, R. (1968) Nature 217, 1110.
2. Smith, H.O., and Wilcox, K.W. (1970) J. Mol. Biol. 51, 379.
3. Roberts, R.J. (1989) Nucleic Acids Res. 17, r347.
4. Yuan, R. (1981) Annu. Rev. Biochem. 50, 285.
5. Petrušytė, M., Bitinaitė, J., Menkevičius, S., Klimašauskas, S., Butkus, V., and Janulaitis, A. (1988) Gene 74, 89.
6. Szybalski, W. (1985) Gene 40, 169.
7. Kessler, C., Neumaier, P.S., and Wolf, W. (1985) Gene 33, 1.
8. Ehrlich, M., Gama-Sosa, M.A., Carreira, L.H., Ljungdahl, L.G., Kuo, K.C., and Gehrke, C.W. (1985) Nucleic Acids Res. 13, 1399.
9. Nelson, M., and McClelland, M. (1987) Nucleic Acids Res. 15, r219.
10. Nelson, M., and Christ, C., and Schildkraut, I. (1984) Nucleic Acids Res. 12, 5165.
11. Mann, M.B., Rao, R.N., and Smith, H.O. (1978) Gene 3, 97.
12. Wilson, G.G.(1988) Gene 74, 281.
13. Pósfai, J., Bhagwat, A.S., and Roberts, R.J. (1988) Gene 74, 261.
14. Lauster, R., Kriebardis, A., and Guschlbauer, W. (1987) FEBS Letters 220, 167.
15. Loenen, W.A.M., Daniel, A.S., Braymer, H.D., and Murray, N.E. (1987) J. Mol. Biol. 198, 159
16. McClarin, J.A., Frederick, C.A., Wang, B.-C., Greene, P.,

Boyer, H.W., Grable, J., and Rosenberg, J.M. (1986) Science 234, 1526.

17. Chandrasegaran, S., Smith, H.O., Amzel, M.L., and Ysern, X. (1986) Prot. Struct. Funct. Genet. 1, 263.
18. Roberts, R.J. (1983) Nucleic Acids Res. 11, r135.
19. Ito, H., Ike, Y., Ikuta, S., and Itakura, K. (1982) Nucleic Acids Res. 10, 1755.
20. Jay, E., Bambara, R., Padmanabhan, R., and Wu, R. (1974) Nucleic Acids Res. 1, 331.
21. Lui, A.C.P., McBride, B.C., Vovis, G.F. and Smith, M (1979) Nucleic Acids Res. 6, 1.
22. Roberts, R.J. (1984) Nucleic Acids Res. 12, r167.
23. Seeman, N.C., Rosenberg, J.M., and Rich, A. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 804.
24. Connolly, B.A., Potter, B.V.L., Eckstein, F., Pingoud, A., and Grotjahn, L. (1984) Biochemistry 23, 3443.
25. Ono, A., Sato, M., Ohtani, Y., and Ueda, T. (1984) Nucleic Acids Res. 12, 8939.
26. Greene, P.J., Poonian, M.S., Nussbaum, A.L., Tobias, L., Garfin, D.E., Boyer, H.W., and Goodman, H.M. (1975) J. Mol. Biol. 99, 237.
27. Goppelt, M., Pingoud, A., Maass, G., Mayer, H., Köster,H., and Frank, R. (1980) Eur. J. Biochem. 104, 101.
28. Thomas, M. and Davis, R.W. (1975) J. Mol. Biol. 91, 315.
29. Guerry, P., LeBlanc, D.J., and Falkow, S. (1973) J. Bacteriol. 116, 1064.

30. Nakata, A., Yamaguchi, M., Izutani, K., and Amemura, M. (1982) J. Bacteriol. 134, 287.

31. Panet, A., van de Sande, J.H., Loewen, P.C., Khorana, H.G. Raae, A.J., Lillehaug, J.R., and Kleppe, K. (1973) Biochemistry 12, 5045.

32. Takahashi, H., Kojima, H., and Saito, H. (1985) Biochem. J. 231, 229.

33. Maxam, A.M., and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560.

34. McDonell, M.W., Simon, M.N., and Studier, F.W. (1977) J. Mol. Biol. 110, 119.

35. Davis, R.W., Botstein, D., and Roth, J.R. (1980) in Advanced Bacterial Genetics, (Davis, R.W. Botstein, D., and Roth, J.R. eds) 222, (Cold Spring Harbor Laboratory), New York.

36. Dwyer-Hallquist, P., Kézdy, F.J., and Agarwal, K.L. (1982) Biochemistry 21, 4693.

37. Wang, A.H.-J., Fujii, S., van Boom, J.H., and Rich, A. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 3968.

38. Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W.B.T., Salisbury, S.A., and Viswamitra, M.A. (1983) J. Mol. Biol. 166, 183.

39. Calladine, C.R. (1982) J. Mol. Biol. 161, 343.

40. Dickerson, R.E. (1983) J. Mol. Biol. 166, 419.

41. Roberts, R.J. (1985) Nucleic Acids Res. 13, r165.

42. Ueda, K., Morita, J., and Komano, T. (1981) J. Antibiot. 34, 317.

43. Greene, P.J., Betlach, M.C., Goodman, H.M., and Boyer, H.W. (1974) Methods in Molecular Biology 7, 87.

44. Old, R., Murray, K., and Roizes, G. (1975) J. Mol. Biol. 92, 331.

45. Walder, R.Y., Hartley, J.L., Donelson, J.E., and Walder, J.A. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 1503.

46. Arrand, J.R., Myers, P.A., and Roberts, R.J. (1978) J. Mol. Biol. 118, 127.

47. Landy, A., Ruedisueli, E., Robinson, L. Foeller, C., and Ross, W. (1974) Biochemistry 13, 2134.

48. Hiraoka, N., Kita, K., Nakajima, H., and Obayashi, A. (1984) J. Ferment. Technol. 62, 583.

49. Geier, G.E., and Modrich, P. (1979) J. Biol. Chem. 254, 1408.

50. Mayer, H., Grosschedl, R., Schütte, H., and Hobom, G. (1981) Nucleic Acids Res. 9, 4833.

51. Sutcliffe, J.G. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 77.

52. Peden, K.W.C. (1983) Gene 22, 277.

53. Sanger, F., Coulson, A.R., Hong, G.F., and Petersen, G.B. (1982) J. Mol. Biol. 162, 729.

54. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., van de Voorde, A., van Heuverswyn, H., van Herreweghe, J., Volckaert, G., and Ysebaert, M. (1978) Nature 273, 113.

55. Sanger, F., Coulson, A.R., Friedman, T., Air, G.M., Barrell, B.G., Brown, N.L., Fiddes, J.C., Hutchison III, C.A., Slocombe, P.M., and Smith, M. (1978) J. Mol. Biol. 125, 225.

56. EMBL nucleotide sequence data library, release 5.0 (April, 1985).

57. Lautop, H. (1974) in Bergey's Manual of Determinative Bacteriology Eighth edition (Buchanan, R.E., and Gibbons, N.E. eds) 436, (The Williams and Wilkins Company), Baltimore.

58. Greene, P.J., Gupta, M., Boyer, H.W., Brown, W.E., and Rosenberg, J.M. (1981) J. Biol. Chem. 256, 2143.

59. Newman, A.K., Rubin, R.A., Kim, S-H., and Modrich, P. (1981) J. Biol. Chem. 256, 2131.

60. Schoner, B., Kelly, S., and Smith, H.O. (1983) Gene 24, 227.

61. Bougueleret, L., Schwarzstein, M., Tsugita, A., and Zabeau, M.(1984) Nucleic Acids Res. 12, 3659.

62. Walder, R.Y., Walder, J.A., and Donelson, J.E. (1984) J. Biol. Chem. 259, 8015.

63. Kiss, A., Pósfai, G., Keller, C.C., Venetianer, P., and Roberts, R.J. (1985) Nucleic Acids Res. 13, 6403.

64. Theriault, G., Roy, P.H., Howard, K.A., Benner, J.S., Brooks, J.E., Waters, A.F., and Gingeras, T.R. (1985) Nucleic Acids Res 13, 8441.

65. Slatko, B.E., Benner, J.S., Jager-Quinton, T., Moran, L.S., Simcox, T.G., Van Cott, E.M., and Wilson, G.G. (1987) Nucleic Acids Res. 15, 9781.

66. Sznyter, L.A., Slatko, B., Moran, L., O'Donnell, K.H., and Brooks, J.E. (1987) Nucleic Acids Res. 15, 8249.

67. Nwankwo, D., and Wilson, G. (1987) Mol. Gen. Genet. 209, 570.

68. Sugisaki, H., and Kanazawa, S. (1981) Gene 16, 73.

69. Boyer, H.W., and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459.

70. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33, 103.

71. Hohn, B., and Collins, J. (1980) Gene 11, 291.

72. Thomas, C.A. Jr., Berns, K.I., and Kelly, T.J. Jr. (1966) in Procedures in Nucleic Acid Research (Cantoni, G.L., and Davies, D.R., eds) 535. (Harper and Row), New York.

73. Birnboim, H.C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513.

74. Maniatis, T. Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harvor Lab.) Cold Spring Harbor, New York.

75. Matsudaira, P. (1987) J. Biol. Chem. 262, 10035.

76. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463.

77. Messing, J. (1983) Methods in Enzymology 101, 20.

78. Brooks, J.E., Blumenthal, R.M., and Gingeras, T.R. (1983) Nucleic Acids Res. 11, 837.

79. Macdonald, P., and Mosig, G.(1984) EMBO J. 3, 2863.

80. Mannarelli, B.M., Balganesh, T.S., Greenberg, B., Springhorn, S.S., and Lacks, S.A. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 4468.

81. Narva, K.E., Wendell, D.L., Skrdla, M., and Van Etten, J.L. (1987) Nucleic Acids Res. 15, 9807.

82. Pósfai, G., Kiss, A., Erdei, S., Pósfai, J., and Venetianer,

P. (1983) J. Mol. Biol. 170, 597.

83. Caserta, M., Zacharias, W., Nwankwo, D., Wilson, G.G., and Wells, R.D. (1987) J. Biol. Chem. 262, 4770.

84. Som, S., Bhagwat, A.S., and Friedman, S. (1987) Nucleic Acids Res. 15, 313.

85. Vermersch, P.S., and Bennett, G.N.(1987) Gene 54, 229.

86. Mandecki, W., and Bolling, T.J. (1988) Gene 68, 101.

87. Kim, S.C., Podhajska, A.J., and Szybalski, W. (1988) Science 240, 504.

88. Ishino, Y., Shinagawa, H., Makino, K., Tsunasawa, S., Sakiyama, F., and Nakata, A. (1986) Mol. Gen. Genet. 204, 1.

89. Chang, A.C.Y., and Cohen, S.N. (1978) J. Bacteriol. 134, 1141.

90. Rose, R.E. (1988) Nucleic Acids Res. 16, 355.

91. Bradford, M.M. (1976) Anal. Biochem. 72, 248.

92. Laemmli, U.K. (1970) Nature 227, 680.

93. Bougueleret, L., Tenchini, M.L., Botterman, J., and Zabeau, M. (1985) Nucleic Acids Res. 13, 3823.

94. Barany, F. (1988) Gene 65, 167.

95. Aiken, C., and Gumport, R.I. (1988) Nucleic Acids Res. 16, 7901.

96. Rubin, R.A., and Modrich, P. (1977) J. Biol. Chem. 252, 7265.

97. Sato, H., and Yamada, Y. (1988) Abstracts of Papers, 110th Meeting of the Agricultural Chemical Society of Japan, Chubu Branch, Gifu, October, Abstract No. 10 (in Japanese).

98. Tsunasawa, S., Stewart, J.W., and Sherman, F. (1985) J. Biol.

Chem. 260, 5382.

99. Rosenberg, J.M., Boyer, H.W., and Greene, P.J. (1981) in Gene Amplification and Analysis (Chirikjian, J.G., ed) vol. 1, 131. Elsevier, N.Y.

100. Chen, S.-C., Kim, R., King, K., Kim, S.-H., and Modrich, P. (1984) J. Biol. Chem. 259, 11571.

101. Botterman, J., and Zabeau, M. (1985) Gene 37, 229.

102. Kelly, S., Kaddurah-Daouk, R., and Smith, H.O. (1985) J. Biol. Chem. 260, 15339.

103. D'Arcy, A., Brown, R.S., Zabeau, M., van Resandt, R.W., and Winkler, F.K. (1985) J. Biol. Chem. 260, 1987..

104. Pósfai, G., and Szybalski, W. (1988) Nucleic Acids Res. 16, 6245.

105. Pósfai, G., and Szybalski, W. (1988) Gene 69, 147.

106. Nishimura, S. (1972) in Progress in Nucleic Acids Research and Molecular Biology (Davidson, J.N., and Cohn, W.E., eds) 49. (Academic Press), New York.

107. Landry, D., Looney, M.C., Feehery, G.R., Slatko, B.E., Jack, W.E., Schildkraut, I., and Wilson, G.G. (1989) Gene, 77, 1.

108. Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987) Methods in Enzymology 154, 367.

109. Chan, P.T., Ohmori, H., Tomizawa, J., and Lebowitz, J. (1985) J. Biol. Chem. 260, 8925.

110. Podhajska, A.J., and Szybalski, W. (1985) Gene 40, 175.

111. Roberts, R.J. (1988) Nucleic Acids Res. 16, r271.

PUBLICATIONS

CHAPTER I

Kita, K., Hiraoka, N., Kimizuka, F., and Obayashi, A. (1984) Agric. Biol. Chem. 48, 531-532.

CHAPTER II

Kita, K., Hiraoka, N., Kimizuka, F., Obayashi, A., Kojima, H., Takahashi, H., and Saito, H. (1985) Nucleic Acids Res. 13, 7015-7024.

CHAPTER III

Kita, K., Hiraoka, N., Oshima, A., Kadonishi, S., and Obayashi, A. (1985) Nucleic Acids Res. 13, 8685-8694.

CHAPTER IV

Kita, K., Kotani, H., Sugisaki, H., and Takanami, M. (1989) J. Biol. Chem. 264, 5751-5756.

CHAPTER V

Kita, K., Kotani, H., Hiraoka, N., Nakamura, T., and Yonaha, K., submitted.

CHAPTER VI

Kita, K., Hiraoka, N., and Nakamura, T., submitted.

CHAPTER VII

Sugisaki, H., Kita, K., and Takanami, M. (1989) J. Biol. Chem. 264, 5757-5761.